

86208

Access DB# _____

SEARCH REQUEST FORM

CRFE

Scientific and Technical Information Center

Requester's Full Name: David Lukton Examiner #: 71263 Date: 2/7/03
Art Unit: 1653 Phone Number 301-3213 Serial Number: 09/380738
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL
Mail Box: 9B01; Exr Rm: 9B05

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Calcium Phosphopeptide Complexes

Inventors (please provide full names):
Reynolds, Eric C.

Earliest Priority Filing Date: 03-13-98

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search Seq ID Nos: 1-5

286 1 AA @ 21
2 @ 25
3 @ 25
4 @ 21
5 @ 54

Edward M. H.
Technical Info. Specialist
STN. Search
GPO: 1997-1-100-000

STAFF USE ONLY

Type of Search

Vendors and cost where applicable

Searcher: _____	NA Sequence (#) _____	STN _____
Searcher Phone #: _____	AA Sequence (#) <u>5</u>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>2/11/03</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>2/10/03</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <u>02</u>
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: _____	Other _____	Other (specify) _____

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:06:55 ; Search time 30.5258 seconds
(without alignments)
91.669 Million cell updates/sec

Title: US-09-380-738A-1
Perfect score: 83
Sequence: 1 QMEAEIXXXEIVPNVBEQK 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	94.0	21	13	Anticariogenic pho
2	78	94.0	21	14	Casein phosphopept
3	78	94.0	21	15	Sequence of casein
4	78	94.0	21	19	Bos alpha-s1-casei
5	78	94.0	37	13	Anticariogenic pho
6	78	94.0	37	15	Casein phosphopept
7	78	94.0	37	15	Sequence of casein
8	78	94.0	192	16	Alpha-s1-casein la
9	75	90.4	21	8	Phosphopeptide 1.
10	75	90.4	21	14	Phosphopeptide 1.

11	75	90.4	21	14	AA323308	Phosphopeptide der
12	75	90.4	21	16	AA689936	Sodium caseinate t
13	73	88.0	21	13	AA28429	Anticariogenic pho
14	73	88.0	21	14	AA32931	Casein phosphopept
15	73	88.0	21	15	AA47818	Sequence of casein
16	73	88.0	21	15	AA12800	Phosphopeptide T1.
17	72	86.7	21	12	AA14447	Phosphopeptide #1.
18	70	84.3	21	13	AA28430	Anticariogenic pho
19	70	84.3	21	14	AA32932	Casein phosphopept
20	70	84.3	21	15	AA47819	Sequence of casein
21	63.5	76.5	20	23	AB77538	Casein with Phe re
22	63.5	76.5	205	16	AA64163	Phosphopeptide S
23	40	48.2	9	19	AAW6605	B. burgdorferi ant
24	39	47.0	299	20	AA19923	B. burgdorferi ant
25	39	47.0	323	20	AA19922	Herbicidally activ
26	39	47.0	1396	23	AB93023	Drosophila melanog
27	38	45.8	466	22	AB58680	Drosophila melanog
28	38	45.8	2421	22	AB65643	Arabidopsis thalia
29	37	44.6	566	21	AA22800	Arabidopsis thalia
30	37	44.6	566	21	AA646524	Arabidopsis thalia
31	37	44.6	579	21	AA22799	Arabidopsis thalia
32	37	44.6	624	21	AA646523	Arabidopsis thalia
33	37	44.6	624	21	AA22798	Arabidopsis thalia
34	37	44.6	624	21	AA646522	Arabidopsis thalia
35	37	44.6	739	22	AAW93207	Human polypeptide,
36	37	44.6	766	21	AA43831	Human cancer assoc
37	37	44.6	766	21	AA43831	Human ovarian anti
38	36.5	44.0	875	23	AB41543	Listeria monocytog
39	36	43.4	113	21	AA615925	Arabidopsis thalia
40	36	43.4	123	22	AB70162	Drosophila melanog
41	36	43.4	130	21	AA613140	Arabidopsis thalia
42	36	43.4	130	21	AA61551	Streptococcus pneu
43	36	43.4	138	21	AA12821	Arabidopsis thalia
44	36	43.4	138	21	AA636403	Arabidopsis thalia
45	36	43.4	139	21	AA636208	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AA28426
ID AAR28426 standard; peptide; 21 AA.
XX AAR28426;
AC AC
XX 19-MAR-1993 (first entry)
DT Anticariogenic phosphopeptide.
XX Casein; metal ion; dietetic; purification; growth medium; dietary
KW supplement; fertiliser.
XX Key Location/Qualifiers
FH Modified-site 1 /note= "may be Glu or pyro Glu"
FT Modified-site 6 /note= "post-translationally phosphorylated serine"
FT Modified-site 8 /note= "post-translationally phosphorylated serine"
FT Modified-site 9 /note= "post-translationally phosphorylated serine"
FT Modified-site 10 /note= "post-translationally phosphorylated serine"
FT Modified-site 17 /note= "post-translationally phosphorylated serine"
FT Modified-site 17 /note= "post-translationally phosphorylated serine"
XX WO9218526-A.
XX 29-OCT-1992.
XX 16-APR-1992; 92WO-AU00175.

21-AUG-1992; 92WO-AU00441.

XX
PT
PT
PT

CC derivative, where the phosphopeptide includes the amino acid sequence:
 CC Ser(P)-Ser(P)-Glu-Glu. The amorphous phases stabilised by the
 CC P and phosphate are a delivery vehicle for co-localisation of Ca,
 CC producing superior anticaries efficacy over prior art. The amorphous
 CC phases stabilised by the phosphopeptides are also useful as dietary
 CC supplements to increase calcium bioavailability and to help prevent
 CC diseases associated with calcium deficiencies. They are particularly
 CC useful for treatment or prevention of dental caries, calcium
 CC malabsorption and bone diseases such as osteoporosis and osteomalacia.
 CC The compositions are useful in humans and in veterinary medicine in
 CC domestic animals such as cattle, sheep, horses and companion animals e.g.
 CC cats and dogs as well as zoo animals. The present sequence represents a
 CC phosphopeptide component of a specifically claimed complex.

XX Sequence 21 AA;

Query Match 94.0%; Score 78; DB 19; Length 21;
 Best Local Similarity 76.2%; Pred. No. 1.7e-08;
 Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
 DB 1 QMEAESISSSEIIVPNSVEQK 21

RESULT 5
 AAR28433
 ID AAR28433 standard; peptide; 37 AA.

XX AC AAR28433;
 XX DT 19-MAR-1993 (first entry)
 XX DE Anticariogenic phosphopeptide.

XX Casein; metal ion; dietetic; purification; growth medium; dietary
 XX supplement; fertiliser.

Key	Location/Qualifiers
Modified-site 4	/note= "post-translationally phosphorylated serine"
Modified-site 6	/note= "post-translationally phosphorylated serine"
Modified-site 22	/note= "post-translationally phosphorylated serine"
Modified-site 24	/note= "post-translationally phosphorylated serine"
Modified-site 25	/note= "post-translationally phosphorylated serine"
Modified-site 26	/note= "post-translationally phosphorylated serine"
Modified-site 33	/note= "post-translationally phosphorylated serine"

WO9218526-A.
 29-OCT-1992.
 16-APR-1992; 92WO-AU00175.
 19-APR-1991; 91AU-0005706.
 (UYME) UNIV MELBOURNE.
 (VICT-) VICTORIAN DAIRY IND AUTHORITY.
 Reynolds EC;
 WPI; 1992-382039/46.

XX Selected phospho-peptide(s) prodn. having anticariogenic
 XX activities etc. - comprises digesting soluble monovalent cation
 XX salt of casein in soln., introducing di- or trivalent metal ion

XX Disclosure; Page 12; 23pp; English.
 XX Ser(P) = post-translationally phosphorylated serine. A mixture of
 CC casein phosphopeptides (CPP) and/or their salts may be used in a
 CC method for treating dental hypersensitivity. Pref. those CPPs
 CC contg. the sequence -Ser(P)-Ser(P)-Ser(P)-predominate. The CPPs
 CC can be extracted from a casein digest.

XX Sequence 21 AA;
 Query Match 94.0%; Score 78; DB 15; Length 21;
 Best Local Similarity 76.2%; Pred. No. 1.7e-08;
 Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
 DB 1 QMEAESISSSEIIVPNSVEQK 21

RESULT 4
 AAW66599
 ID AAW66599 standard; peptide; 21 AA.

XX AC AAW66599;
 XX DT 27-NOV-1998 (first entry)
 XX DE Bos alpha-s1-casein X-5P (f59-79) phosphopeptide.

XX casein; calcium phosphate complex; amorphous calcium phosphate; ACP;
 XX phosphopeptide; delivery vehicle; calcium fluoride; calcium deficiency;
 XX osteoporosis; osteomalacia; tooth; bone disease.

XX Synthetic.
 XX Bos taurus.

Key	Location/Qualifiers
Modified-site 6	/note= "Ser(P)"
Modified-site 8	/note= "Ser(P)"
Modified-site 9	/note= "Ser(P)"
Modified-site 10	/note= "Ser(P)"
Modified-site 17	/note= "Ser(P)"
Modified-site 17	/note= "Ser(P)"

WO9840406-A1.
 17-SEP-1998.
 13-MAR-1998; 98WO-AU00160.
 13-MAR-1997; 97AU-0005662.
 (UYME) UNIV MELBOURNE.
 (VICT-) VICTORIAN DAIRY IND AUTHORITY.
 Reynolds EC;
 WPI; 1998-520803/44.

XX Stable calcium phosphate complex including phospho-peptide
 XX stabilised amorphous calcium phosphate useful for treatment of
 XX dental caries, calcium malabsorption and bone diseases such as
 XX osteoporosis and osteomalacia.

XX Claim 4; Page 35; 43pp; English.
 XX The invention relates to a stable calcium phosphate complex including
 CC phosphopeptide stabilised amorphous calcium phosphate (ACP) or its

PT and filtering through filter having mol. wt. exclusion limit
 PS Claim 11; Page 14-5; 18pp; English.
 XX
 CC The peptide may be prep'd. by completely digesting casein in soln.
 CC with a proteolytic enzyme, adding mineral acid to the soln. to
 CC adjust the pH to 4.7, removing any precipitate, adding calcium
 CC chloride to cause aggregation of the peptides in soln. and
 CC separating the aggregated phosphopeptides. This method allows prodn.
 CC of the phosphopeptide by industrial methods. The phosphopeptide
 CC has anticariogenic activity and may be used as a dietetic. The
 CC peptide may be used in a microbiological growth medium, as a dietary
 CC supplement or as a fertiliser.
 CC See also AAR28425-32.
 XX
 XX

SQ Sequence 37 AA;

Query Match 94.0%; Score 78; DB 13; Length 37;
 Best Local Similarity 76.2%; Pred. No. 3.3e-08;
 Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
 DB 17 QMEAESISSSEIIVNSVEQK 37

RESULT 6

AAR32935
 ID AAR32935 standard; Protein; 37 AA.
 AC AAR32935;

DT 02-JUL-1993 (first entry)

DE Casein phosphopeptide #9.

KW Casein; phosphopeptide; dental calculus; salts; alkaline metal;
 KW alkaline earth metal; Zn/phosphopeptide complex; aggregate;
 KW anti-carries; anti-gingivitis.
 XX Synthetic.

Key	Location/Qualifiers
Modified-site 4	/label= Phosphoserine
Modified-site 6	/label= Phosphoserine
Modified-site 22	/label= Phosphoserine
Modified-site 24	/label= Phosphoserine
Modified-site 25	/label= Phosphoserine
Modified-site 26	/label= Phosphoserine
Modified-site 33	/label= Phosphoserine
Modified-site	/label= Phosphoserine

W09303707-A.

04-MAR-1993.

21-AUG-1992; 92WO-AU00441.

22-AUG-1991; 91US-0748344.

(UYME) UNIV MELBOURNE.
 (VICT-) VICTORIAN DAIRY IND AUTHORITY.

Reynolds EC;

WPI; 1993-093685/11.

XX

PT Controlling dental calculus by treating teeth with oral compsns.
 PT - which contains phospho-peptide(s) having 5-40 amino-acyl
 PT residues

XX Claim 3; Page 21; 23pp; English.

XX The sequences given in AAR32927-35 are casein phosphopeptides which can
 CC be used to inhibit dental calculus. These peptides are pref. in the
 CC form of salts selected from alkaline metal, alkaline earth metal salts
 CC such as Na, Ca, Zn, Cu, Al, K, Sr, Mg and Ni salts. These peptides
 CC are pref. present as a Zn/phosphopeptide complex or aggregate. These
 CC peptides have anti-calculus potential, and are anti-carries and anti-
 CC gingivitis agents.

SQ Sequence 37 AA;

Query Match 94.0%; Score 78; DB 14; Length 37;
 Best Local Similarity 76.2%; Pred. No. 3.3e-08;
 Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
 DB 17 QMEAESISSSEIIVNSVEQK 37

RESULT 7

AAR47822
 ID AAR47822 standard; peptide; 37 AA.
 AC AAR47822;

DT 21-JUL-1994 (first entry)

DE Sequence of casein phosphopeptide (CPP).

KW Casein phosphopeptide; dentinal hypersensitivity; therapy; tooth.
 KW Synthetic.

Key	Location/Qualifiers
Modified-site 4	/label= Ser(P) /note= "see also AAs 6,22,24,25,26,33"

W09400146-A.

06-JAN-1994.

29-JUN-1993; 93WO-AU00319.

29-JUN-1992; 92AU-0003221.

(UYME) UNIV MELBOURNE.
 (VICT-) VICTORIAN DAIRY IND AUTHORITY.

Reynolds EC;

WPI; 1994-025888/03.

Treatment of dentinal hypersensitivity - using casein, component
 of casein, phospho-protein or phospho-peptide or their salts

Disclosure; Page 18; 23pp; English.

XX Ser(P) = post-translationally phosphorylated serine. A mixture of
 CC casein phosphopeptides (CPP) and/or their salts may be used in a
 CC method for treating dentinal hypersensitivity. Pref. those CPPs
 CC contg. the sequence -Ser(P)-Ser(P)-predominate. The CPPs
 CC can be extracted from a casein digest.

SQ Sequence 37 AA;

Query Match 94.0%; Score 78; DB 15; Length 37;

```

Best Local Similarity 76.2%; Pred. No. 3.3e-08;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
Db 17 QMEAESISSSEIIVPNSVEQK 37

RESULT 8
AAR64162
ID AAR64162 standard; Protein; 192 AA.
XX
AC AAR64162;
XX
DT 03-AUG-1995 (first entry)
XX
DE Alpha-s1-casein lacking Phe residues.
XX
KW Alpha-s1-casein; phenylalanine; phenylketonuria; food;
KW Saccharomyces cerevisiae.
XX
OS Synthetic.
XX
PN WO9428126-A.
XX
PD 08-DEC-1994.
XX
PF 16-MAY-1994; 94WO-GB01046.
XX
PR 20-MAY-1993; 93GB-0010472.
XX
PA (UYWA-) UNIV WARWICK.
XX
PI Carr NG, Mann NH;
XX
DR WPI; 1995-022798/03.
XX
N-PSDB; AAQ75411; AAQ75412.
XX
PT DNA encoding a food protein with Phe codon(s) omitted or replaced
PT with codon(s) for other metabolisable amino acids - used to form
PT edible polypeptide for use in nutrition for phenylketonuria
PT patients
XX
PS Disclosure; Fig. 15; 90pp; English.
XX
CC The sequence of cattle alpha-1-casein, lacking phenylalanine
CC residues, and with an additional N-terminal methionine (to
CC permit translation) is given in AAR64162. The encoding gene
CC was designed for optimal expression in Saccharomyces
CC cerevisiae.
XX
SQ Sequence 192 AA;

Query Match 94.0%; Score 78; DB 16; Length 192;
Best Local Similarity 76.2%; Pred. No. 2.3e-07;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
Db 56 QMEAESISSSEIIVPNSVEQK 76

RESULT 9
AAR71320
ID AAR71320 standard; peptide; 21 AA.
XX
AC AAR71320;
XX
DT 07-MAY-1991 (first entry)
XX
DE Phosphopeptide 1.
XX
KW Caries; gingivitis; periodontal disease; osteoporosis; osteomalacia.

Best Local Similarity 76.2%; Pred. No. 6.6e-08;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
Db 1 EMEAESISSSEIIVPNSVEQK 21

RESULT 10
AAR31237
ID AAR31237 standard; peptide; 21 AA.
XX
AC AAR31237;
XX
DT 18-MAY-1993 (first entry)
XX
DE Phosphopeptide 1.
XX
KW Phosphopeptide; active agent; oral; composition; anionic; polymeric;
KW stabiliser; carboxylate; polymer; sulfonate; destabilisation;
KW fluoride; caries; gingivitis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 6 /label= Phosphoserine
FT Modified-site 8 /label= Phosphoserine
FT Modified-site 9 /label= Phosphoserine
FT Modified-site 10 /label= Phosphoserine
FT Modified-site 17 /label= Phosphoserine
FT Modified-site 9

Query Match 90.4%; Score 75; DB 8; Length 21;
Best Local Similarity 71.4%; Pred. No. 6.6e-08;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
Db 1 EMEAESISSSEIIVPNSVEQK 21

RESULT 10
AAR31237
ID AAR31237 standard; peptide; 21 AA.
XX
AC AAR31237;
XX
DT 18-MAY-1993 (first entry)
XX
DE Phosphopeptide 1.
XX
KW Phosphopeptide; active agent; oral; composition; anionic; polymeric;
KW stabiliser; carboxylate; polymer; sulfonate; destabilisation;
KW fluoride; caries; gingivitis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 6 /label= Phosphoserine
FT Modified-site 8 /label= Phosphoserine
FT Modified-site 9 /label= Phosphoserine
FT Modified-site 9

Synthetic.
Key Location/Qualifiers
Modified-site 6 /label= phosphoserine
Modified-site 8 /label= phosphoserine
Modified-site 9 /label= phosphoserine
Modified-site 10 /label= phosphoserine
Modified-site 17 /label= phosphoserine
Modified-site 17 /label= phosphoserine
WO8707616-A.
17-DEC-1987.
12-JUN-1987; 87WO-AU00172.
12-JUN-1986; 86AU-0006385.
(VICT-) VICTORIA DAIRY INDUSTRY AUTHORITY.
(UYME-) UNIVERSITY OF MELBOURNE.
(REYN/) EC REYNOLDS.
Reynolds EC;
WPI; 1987-362707/51.
New phosphopeptides contg. defined amino acid sequence - useful in
treatment of dental, rarefying bone diseases and disease relating to
malabsorption of minerals.
Claim 4; Page 17; 22pp; English.
The phosphopeptide is used in compsns. at a conc. of 0.01-5 wt%.
See also AAP71321-P71324.
Sequence 21 AA;

Query Match 90.4%; Score 75; DB 8; Length 21;
Best Local Similarity 71.4%; Pred. No. 6.6e-08;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
Db 1 EMEAESISSSEIIVPNSVEQK 21

RESULT 10
AAR31237
ID AAR31237 standard; peptide; 21 AA.
XX
AC AAR31237;
XX
DT 18-MAY-1993 (first entry)
XX
DE Phosphopeptide 1.
XX
KW Phosphopeptide; active agent; oral; composition; anionic; polymeric;
KW stabiliser; carboxylate; polymer; sulfonate; destabilisation;
KW fluoride; caries; gingivitis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 6 /label= Phosphoserine
FT Modified-site 8 /label= Phosphoserine
FT Modified-site 9 /label= Phosphoserine
FT Modified-site 9

```

FT Modified-site 10 /label= Phosphoserine
 FT Modified-site 17 /label= Phosphoserine
 FT Modified-site 17 /label= Phosphoserine
 XX EP523776-A.
 XX 20-JAN-1993.
 XX 03-JUL-1992; 92EP-0202023.
 XX 17-JUL-1991; 91US-0731592.
 XX (UNIL) UNILEVER NV.
 XX (UNIL) UNILEVER PLC.
 XX Burger AR, Elliott DL, Schick LA;
 XX WPI; 1993-019802/03.
 XX Oral compns. contg. a phosphopeptide - with addn. of an anionic
 FT polymeric stabiliser to inhibit destabilisation in the oral
 FT environment
 XX Claim 6; Page 13; 18pp; English.
 XX The sequences given in AAR31237-42 represent phosphopeptides which
 CC were used as the active agents in an oral composition. These
 CC peptides were stabilised by an anionic polymeric stabiliser. The
 CC anionic polymers were chosen from a group consisting of carboxylate
 CC polymers, sulfonate polymers, polymers having both a carboxylate and
 CC a sulfonate moiety, and other such mixtures. The anionic polymeric
 CC stabiliser inhibits destabilisation of the phosphopeptide in the oral
 CC environment. These oral compositions, pref. containing a fluoride
 CC source may be used for inhibiting caries and gingivitis.
 XX Sequence 21 AA;
 SQ Query Match 90.4%; Score 75; DB 14; Length 21;
 Best Local Similarity 71.4%; Pred. No. 6.6e-08;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 OMEAEIXXXEIVPNXVEQK 21
 Db :||||| | ||||| |||||
 1 EMEAESISSSEIVPNXVEQK 21
 RESULT 11
 AAR32308
 ID AAR32308 standard; peptide; 21 AA.
 XX
 AC AAR32308;
 XX
 DT 10-JUN-1993 (first entry)
 XX
 DE Phosphopeptide derived from casein.
 XX
 KW Dental; teeth; tartar control; brushite; calcium phosphate;
 KW hydroxyapatite; mouthwash; toothpaste.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 6
 FT /label= "post-translationally phosphorylated"
 FT Modified-site 8
 FT /note= "post-translationally phosphorylated"
 FT Modified-site 9
 FT /note= "post-translationally phosphorylated"
 FT Modified-site 10
 FT /note= "post-translationally phosphorylated"
 FT Modified-site 17
 FT /note= "post-translationally phosphorylated"

FT XX /note= "post-translationally phosphorylated"
 PN EP528458-A.
 XX 24-FEB-1993.
 PD
 XX 03-JUL-1992; 92EP-0202024.
 XX 09-AUG-1991; 91GB-0017315.
 XX (UNIL) UNILEVER PLC.
 PA (UNIL) UNILEVER NV.
 XX Burger AR, Schick LA;
 XX WPI; 1993-060322/08.
 XX Phospho-peptide(s) for dental tartar control - are included in
 FT compns. with pyrophosphate(s) or zinc salts to provide good
 FT inhibition
 XX Claim 3; Page 11; 17pp; English.
 XX The peptide sequence is that of a phosphopeptide prep. from a
 CC tryptic digest of casein. The peptide may be used with an anti-
 CC calculus agent, e.g. alkali metal pyrophosphate, in a compsn. for
 CC controlling dental tartar. The compsn. inhibits conversion of
 CC brushite and amorphous calcium phosphate into the more stable
 CC hydroxyapatite on the teeth. The compsn. is used in the form of a
 CC mouthwash, toothpaste, gel, lozenge or chewing gum, for care of the
 CC teeth and tartar control.
 CC See also AAR32309-12.
 XX Sequence 21 AA;
 SQ Query Match 90.4%; Score 75; DB 14; Length 21;
 Best Local Similarity 71.4%; Pred. No. 6.6e-08;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 OMEAEIXXXEIVPNXVEQK 21
 Db :||||| | ||||| |||||
 1 EMEAESISSSEIVPNXVEQK 21
 RESULT 12
 AAR68936
 ID AAR68936 standard; peptide; 21 AA.
 XX
 AC AAR68936;
 XX
 DT 07-SEP-1995 (first entry)
 XX
 DE Sodium caseinate tryptic phosphopeptide Tl.
 XX
 KW Phosphopeptide; tryptic digestion; sodium caseinate; alpha(sI)-caseinate;
 KW phosphoserine; phosphothreonine; phosphotyrosine; phosphohistidine;
 KW sugar; sorbitol; mannitol; xylitol; lactitol; cellobitol; confectionary;
 KW caries; gingivitis; calcium; remineralisation.
 XX
 OS Bos taurus.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 6
 FT /label= OTHER
 FT Modified-site 8
 FT /note= "Phosphoserine"
 FT Modified-site 9
 FT /label= OTHER
 FT Modified-site 9
 FT /note= "Phosphoserine"
 FT Modified-site 10
 FT /label= OTHER
 FT Modified-site 10
 FT /note= "Phosphoserine"
 FT Modified-site 10
 FT /label= OTHER


```

FT XX                               /label= Phosphoserine
PN XX
PD WO9303707-A.
XX
XX 04-MAR-1993.
XX
XX 21-AUG-1992; 92WO-AU00441.
XX
XX 22-AUG-1991; 91US-0748344.
XX
XX (UYME ) UNIV MELBOURNE.
PA (VICT-) VICTORIAN DAIRY IND AUTHORITY.
XX
XX Reynolds EC;
XX
XX WPI; 1993-093685/11.
XX
XX Controlling dental calculus by treating teeth with oral compens.
PT - which contains phospho-peptide(s) having 5-40 amino-acyl
PT residues
XX
XX Claim 3; Page 18; 23pp; English.
XX
XX The sequences given in AAR32927-35 are casein phosphopeptides which can
CC be used to inhibit dental calculus. These peptides are pref. in the
CC form of salts selected from alkaline metal alkaline earth metal salts
CC such as Na, Ca, Zn, Cu, Al, K, Sr, Mg and Ni salts. These peptides
CC are pref. present as a Zn/phosphopeptide complex or aggregate. These
CC peptides have anti-calculus potential, and are anti-carries and anti-
CC gingivitis agents.
XX
XX Sequence 21 AA;
XX
Query Match 88.0%; Score 73; DB 14; Length 21;
Best Local Similarity 71.4%; Pred. No. 1.6e-07;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 QMEAEIXXXEIEIVPNXVEQK 21
Db 1 QMEAESISSSEIEIVPDSVEQK 21
RESULT 15
AAR47818
ID AAR47818 standard; peptide; 21 AA.
AC AAR47818;
XX
XX 21-JUL-1994 (first entry)
XX
XX Sequence of casein phosphopeptide (CPP).
XX
XX Casein phosphopeptide; dental hypersensitivity; therapy; tooth.
XX Synthetic.
XX
XX Key Location/Qualifiers
FH 6
FT Modified-site
FT /label= Ser(P)
FT /note= "see also residues 8,9,10,17"
XX
XX WO9400146-A.
XX
XX 06-JAN-1994.
XX
XX 29-JUN-1993; 93WO-AU00319.
XX
XX 29-JUN-1992; 92AU-0003221.
XX
XX (UYME ) UNIV MELBOURNE.
PA (VICT-) VICTORIAN DAIRY IND AUTHORITY.
XX
XX Reynolds EC;
PI

```

```

XX
XX WPI; 1994-025888/03.
XX
XX Treatment of dental hypersensitivity - using casein, component
PT of casein, phospho-protein or phospho-peptide or their salts
XX
XX Disclosure; Page 14; 23pp; English.
XX
XX Ser(P) = post-translationally phosphorylated serine. A mixture of
CC casein phosphopeptides (CPP) and/or their salts may be used in a
CC method for treating dental hypersensitivity. Pref. those CPPs
CC contg. the sequence -Ser(P)-Ser(P)-Ser(P)- predominate. The CPPs
CC can be extracted from a casein digest.
XX
XX Sequence 21 AA;
XX
Query Match 88.0%; Score 73; DB 15; Length 21;
Best Local Similarity 71.4%; Pred. No. 1.6e-07;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 QMEAEIXXXEIEIVPNXVEQK 21
Db 1 QMEAESISSSEIEIVPDSVEQK 21
Search completed: February 11, 2003, 18:16:53
Job time : 32.5258 secs

```

us-09-380-738a-1.rai

Wed Feb-12 11:35:20 2003

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
February 11, 2003, 18:14:31 : Search time 10.3918 seconds
(without alignments)
59.459 Million cell updates/sec

US-09-380-738A-1
Title: 83
Perfect score: 1 QMEAXIXXEEIVPNXVEOK 21
Sequence:

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues 262574

Total number of hits satisfying chosen parameters:
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS-COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	94.0	21	1	US-07-748-344B-2
2	78	94.0	21	2	US-08-954-985A-2
3	78	94.0	21	4	US-08-137-086-2
4	78	94.0	37	1	US-07-748-344B-9
5	78	94.0	37	2	US-08-954-985A-9
6	78	94.0	37	1	US-08-137-086-9
7	73	88.0	21	1	US-07-748-344B-5
8	73	88.0	21	2	US-08-954-985A-5
9	70	84.3	21	1	US-07-748-344B-6
10	70	84.3	21	2	US-08-954-985A-6
11	70	84.3	21	4	US-08-137-086-6
12	70	84.3	20	4	US-09-443-339A-1
13	63.5	76.5	1311	1	US-08-340-011-5
14	36	43.4	1311	3	US-08-901-710-5
15	36	43.4	1338	3	US-08-750-141A-3
16	36	43.4	803	1	US-08-062-368-4
17	35	42.2	803	2	US-08-907-166-4
18	34	41.0	803	1	US-08-072-610-2
19	34	41.0	1018	2	US-08-719-822B-2
20	34	41.0	1018	4	US-09-092-458-2
21	34	41.0	1018	4	US-09-021-560-4
22	33	39.8	707	4	US-09-134-001C-4037
23	33	39.8	1076	4	US-08-769-309A-5
24	33	39.8	1780	1	US-08-994-570-5
25	33	39.8	125	4	US-08-975-762-51
26	32	38.6	125	4	US-09-295-028-51
27	32	38.6	125	4	US-09-295-028-51

Sequence 51, Appl
Sequence 2, Appl
Sequence 13, Appl
Sequence 23, Appl
Sequence 23, Appl
Sequence 23, Appl
Sequence 23, Appl
Sequence 5, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 6, Appl
Sequence 50, Appl
Sequence 50, Appl
Sequence 8, Appl
Sequence 45, Appl
Sequence 8, Appl
US-09-106-582-51
US-09-068-140A-2
US-09-068-140A-13
US-08-975-762-23
US-08-821-324-23
US-08-285-028-23
US-09-106-582-23
US-09-106-888A-5
US-08-606-715A-12
US-08-451-635A-12
US-08-500-635A-12
US-09-167-151-12
US-09-066-046-6
US-08-975-762-50
US-09-295-028-50
US-09-106-582-50
US-08-973-462-8
US-09-091-725-45
US-08-506-553C-8

ALIGNMENTS

RESULT 1
US-07-748-344B-2 Application US/07748344B
Sequence 2, Appl
Patent No. 5227154
GENERAL INFORMATION: REYNOLDS, ERIC CHARLES
APPLICANT: REYNOLDS, ERIC CHARLES
TITLE OF INVENTION: PHOSPHOTETIDES FOR THE
TREATMENT OF DENTAL CALCULUS
TITLE OF INVENTION: PHOSPHOTETIDES FOR THE
TREATMENT OF DENTAL CALCULUS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: ANDRUS, SCEALES, STARKE & SAWALL
STREET: 100 EAST WISCONSIN AVE., SUITE 1100
CITY: MILWAUKEE
STATE: WISCONSIN
COUNTRY: USA
ZIP: 53202

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07748, 344B
FILING DATE: 19910822
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: SARA, CHARLES S
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: C.8493-87
TELEPHONE: (608) 255-2022
TELEFAX: (608) 255-2182
TELEX: 26832 ANDSTARK
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 21
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
FEATURE:
NAME/KEY: Pyroglutamate
LOCATION: 1
OTHER INFORMATION: A certain amount will exist in this form
OTHER INFORMATION:
NAME/KEY: Phosphoserine

us-09-380-738a-1.rai

```

; LOCATION: 6
; OTHER INFORMATION:
; FEATURE: Post-translationally phosphorylated serine
; NAME/KEY: Phosphoserine
; LOCATION: 8
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 9
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 10
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 17
; OTHER INFORMATION:
; US-07-748-344B-2

```

Query Match 94.0%; Score 78; DB 1; Length 21;

Best Local Similarity 76.2%; Pred. No. 9e-09; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 QMEAEIXXXEIVPNXVEQK 21
Db 1 QMEAESISSSEIVPNXVEQK 21

```

RESULT 2

US-08-954-985A-2
Sequence 2, Application US/08954985A
Patent No. 5981475
GENERAL INFORMATION:

APPLICANT: REYNOLDS, Eric Charles
TITLE OF INVENTION: A treatment for Sensitive Teeth
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

ADDRESSER: Dental School, The University of Melbourne
CITY: Melbourne
STATE: Victoria
COUNTRY: Australia
ZIP: 3000

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM
OPERATING SYSTEM: Windows 95

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/954,985A
FILING DATE: 21-OCT-1997
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,479
FILING DATE: 22-FEB-1995
ATTORNEY/AGENT INFORMATION:

NAME: WOZNY, Thomas M
REGISTRATION NUMBER: 28,922
REFERENCE/DOCKET NUMBER: 322-00033
TELECOMMUNICATION INFORMATION:

TELEPHONE: (414) 271-5770
TELEFAX: (414) 271-5770
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 21
TYPE: Amino Acid
TOPOLOGY: Linear

```

; FEATURE:
; NAME/KEY: Pyroglutamate
; LOCATION: 1
; OTHER INFORMATION:
; FEATURE: A certain amount will exist in this form
; NAME/KEY: Phosphoserine
; LOCATION: 6
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 8
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 9
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 10
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 17
; OTHER INFORMATION:
; US-08-954-985A-2

```

Query Match 94.0%; Score 78; DB 2; Length 21;

Best Local Similarity 76.2%; Pred. No. 9e-09; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 QMEAEIXXXEIVPNXVEQK 21
Db 1 QMEAESISSSEIVPNXVEQK 21

```

RESULT 3

US-08-137-086-2
Sequence 2, Application US/08137086
Patent No. 6448374
GENERAL INFORMATION:

APPLICANT: REYNOLDS, ERIC CHARLES
TITLE OF INVENTION: PRODUCTION OF PHOSPHOPEPTIDES FROM CASEIN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

ADDRESSEE: ANDRUS, SCALES, STARKE & SAWALL
STREET: 100 EAST WISCONSIN AVE., SUITE 1100
CITY: MILWAUKEE
STATE: WISCONSIN
COUNTRY: USA
ZIP: 53202

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,086
FILING DATE: 04-MAR-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: SARA, CHARLES S

REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: C-8493-87
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 255-2022
TELEFAX: (608) 255-2182

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

us-09-380-738a-1.rai

Wed Feb-12 11:35:20 2003

```

;
; LENGTH: 21
; TYPE: Amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; FEATURE:
; NAME/KEY: Pyroglutamate
; LOCATION: 1
; OTHER INFORMATION: A certain amount will exist in this form
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 6
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 8
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 9
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 10
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 17
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; US-08-137-086-2
;
; Query Match 94.0%; Score 78; DB 4; Length 21;
; Best Local Similarity 76.2%; Pred. No. 9e-09;
; Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
;
; QY 1 QMEAEIXXXEIEIVPNVQK 21
;      ||||| | ||||| |||||
; Db 1 QMEAESISSSEIEIVPNVQK 21
;
; RESULT 4
; US-07-748-344B-9
; Sequence 9, Application US/07748344B
; Patent No. 5227154
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; TITLE OF INVENTION: PHOSPHOPEPTIDES FOR THE
; TITLE OF INVENTION: TREATMENT OF DENTAL CALCULUS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDRUS, SCALES, STARKE & SAWALL
; STREET: 100 EAST WISCONSIN AVE., SUITE 1100
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07748.344B
; FILING DATE: 19910822
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
;
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: SARA, CHARLES S
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: C.8493-87
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 255-2022
; TELEFAX: (608) 255-2182
; TELEX: 26832 ANDSTARK
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 4
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 6
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 22
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 24
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 25
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 26
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 33
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; US-07-748-344B-9
;
; Query Match 94.0%; Score 78; DB 1; Length 37;
; Best Local Similarity 76.2%; Pred. No. 1.7e-08;
; Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
;
; QY 1 QMEAEIXXXEIEIVPNVQK 21
;      ||||| | ||||| |||||
; Db 17 QMEAESISSSEIEIVPNVQK 37
;
; RESULT 5
; US-08-954-985A-9
; Sequence 9, Application US/08954985A
; Patent No. 5981475
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; TITLE OF INVENTION: A Treatment for Sensitive Teeth
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dental School, The University of Melbourne
; STREET: 711 Elizabeth Street
;

```

```

; CITY: Melbourne
; STATE: Victoria
; COUNTRY: Australia
; ZIP: 3000
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM
; OPERATING SYSTEM: Windows 95
; SOFTWARE: ASCII
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954.985A
; FILING DATE: 21-OCT-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,479
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Wozny, Thomas M
; REGISTRATION NUMBER: 28,922
; REFERENCE/DOCKET NUMBER: 322-00033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 271-7590
; TELEFAX: (414) 271-5770
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37
; TYPE: Amino Acid
; TOPOLOGY: Linear
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 4
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 6
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 22
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 25
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 26
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 33
; OTHER INFORMATION: Post-translationally phosphorylated serine
;
; US-08-954-985A-9
;
; Query Match 94.0%; Score 78; DB 2; Length 37;
; Best Local Similarity 76.2%; Pred. No. 1.7e-08;
; Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
;
; QY 1 OMEAEIXXXEIEIVPNXVEQK 21
; Db 17 OMEAESISSEIEIVPNVBEQK 37
;
; RESULT 6
; US-08-137-086-9
; Sequence 9, Application US/08137086
; Patent No. 6448374
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES

```

```

; TITLE OF INVENTION: PRODUCTION OF PHOSPHOPEPTIDES FROM CASEIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDRUS, SCEALES, STARKE & SAWALL
; STREET: 100 EAST WISCONSIN AVE., SUITE 1100
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,086
; FILING DATE: 04-MAR-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: SARA, CHARLES S
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: C.8493-87
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 255-2022
; TELEFAX: (608) 255-2182
; TELEX: 26832 ANDSTARK
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37
; TYPE: Amino acid
; STRANDEDNESS: single
; TOPOLOGY: Linear
; MOLECULE TYPE: Protein
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 4
; OTHER INFORMATION:
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 6
; OTHER INFORMATION:
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 22
; OTHER INFORMATION:
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 24
; OTHER INFORMATION:
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 25
; OTHER INFORMATION:
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 26
; OTHER INFORMATION:
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 33
; OTHER INFORMATION:
; OTHER INFORMATION: Post-translationally phosphorylated serine
;
; US-08-137-086-9

```



```

; LOCATION: 10
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
;
; NAME/KEY: Phosphoserine
; LOCATION: 17
; OTHER INFORMATION:
;
; OTHER INFORMATION: Post-translationally phosphorylated serine
; US-08-137-086-5

Query Match 88.0%; Score 73; DB 4; Length 21;
Best Local Similarity 71.4%; Pred. No. 8.2e-08;
Matches 15; Conservative 1; Mismatches 5; Indels 0;

QY 1 QMEAEXXXXXERIPNVXVEQK 21
      ||||| | ||||| |||||
Db 1 QMEAESISSSEETVPDSVEQK 21

RESULT 10
US-07-748-344B-6
; Sequence 6, Application US/07748344B
; Patent No. 5227154
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; TITLE OF INVENTION: PHOSPHOPEPTIDES FOR THE
; TITLE OF INVENTION: TREATMENT OF DENTAL CALCULUS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDRUS, SCEALES, STARKE & SAWALL
; STREET: 100 EAST WISCONSIN AVE., SUITE 1100
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/748,344B
; FILING DATE: 19910822
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: SARA, CHARLES S
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: C.8493-87
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 255-2022
; TELEFAX: (608) 255-2182
; TELE: 26832 ANDSTARK
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 6
; OTHER INFORMATION:
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE: Phosphoserine
; FEATURE: 8
; FEATURE:
; FEATURE: Post-translationally phosphorylated serine
; FEATURE:

```

us-09-380-738a-1.rai

wed Feb '12 11:35:20 2003

```

; NAME/KEY: phosphoserine
; LOCATION: 9
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 10
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 17
; OTHER INFORMATION: Post-translationally phosphorylated serine
; US-08-954-985A-6
Query Match 84.3%; Score 70; DB 2; Length 21;
Best Local Similarity 66.7%; Pred. No. 3.le-07;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIEIVPNXVEQK 21
    ||||| ||||| ||||| |||||
Db 1 QMEAESISSSEIIVPDSVEEK 21

RESULT 12
US-08-137-086-6
; Sequence 6, Application US/08137086
; Patent No. 6448374
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; TITLE OF INVENTION: PRODUCTION OF PHOSPHOPEPTIDES FROM CASEIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDRUS, SCEALES, STARKE & SAWALL
; STREET: 100 EAST WISCONSIN AVE., SUITE 1100
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,086
; FILING DATE: 04-MAR-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: SARA, CHARLES S
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: C.8493-87
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 255-2022
; TELEFAX: (608) 255-2182
; TELEX: 26832 ANDSTARK
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: Amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 6
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 8
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1
; US-09-443-339A-1
Query Match 76.5%; Score 63.5
Best Local Similarity 95.2%; Pred. No. 1
Matches 20; Conservative 0; Mismatch
;
QY 1 QMEAEAXXXEEIVPNXVBQK 21
Db 1 QMEAEAXI-XXEEIVPNXVBQK 20
;
RESULT 14
US-08-340-011-5
; Sequence 5, Application US/08340011
; Patent No. 5776755
; GENERAL INFORMATION:
; APPLICANT: Alitalo, et al.
; TITLE OF INVENTION: FIT4, A NOVEL RECEPTOR
; NUMBER OF SEQUENCES: 20
; CORRESPONDING ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, et al.
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version 1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,011
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/959,951
; FILING DATE: 09-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Thomas C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: 32267
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-340-011-5
Query Match 43.4%; Score 36.1
Best Local Similarity 46.7%; Pred. No. 1
Matches 7; Conservative 2; Mismatch
;
QY 2 MEAEAXXXEEIVPN 16
Db 1258 MSLEIKTFEELPN 1272
;
RESULT 15
US-08-901-710-5
; Sequence 5, Application US/08901710
; Patent No. 6107046
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:19:51 ; Search time 10.1753 Seconds
(without alignments)
52.729 Million cell updates/sec

Title: US-09-380-738A-1

Perfect score: 83

Sequence: 1 OMEAEIXXXEIVPNXVEQK 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	37	44.6	766	10	US-09-925-301-1276 Sequence 1276, Ap
2	35	42.2	253	10	US-09-864-761-37733 Sequence 37733, A
3	34	41.0	374	10	US-09-925-297-524 Sequence 524, App
4	34	41.0	803	9	US-10-029-382-2 Sequence 2, Appli
5	34	41.0	803	10	US-09-391-340-4 Sequence 4, Appli
6	34	41.0	803	10	US-09-948-369-4 Sequence 14, Appl
7	34	41.0	803	10	US-09-948-369-14 Sequence 16, Appl
8	34	41.0	803	10	US-09-948-369-16 Sequence 2, Appli
9	34	41.0	803	12	US-10-034-849-2 Sequence 2, Appli
10	34	41.0	803	12	US-10-034-621-2 Sequence 5358, Ap
11	34	41.0	1213	10	US-09-815-242-5358 Sequence 12523, A
12	34	41.0	1217	10	US-09-815-242-12523 Sequence 10892, A
13	33	39.8	305	10	US-09-815-242-10892 Sequence 2, Appli
14	33	39.8	708	8	US-08-834-666A-2 Sequence 3, Appli
15	33	39.8	1781	10	US-09-738-877-3 Sequence 51, Appl
16	32	38.6	125	10	US-09-159-469-51 Sequence 51, Appl
17	32	38.6	125	10	US-09-798-042-51 Sequence 51, Appl
18	32	38.6	162	10	US-09-800-528-2 Sequence 2, Appli
19	32	38.6	173	10	US-09-800-528-13 Sequence 13, Appli

20 32 38.6 233 10 US-09-159-469-23 Sequence 23, Appli
21 32 38.6 233 10 US-09-798-042-23 Sequence 23, Appli
22 32 38.6 303 10 US-09-734-569-164 Sequence 164, App
23 32 38.6 424 9 US-10-046-961-7 Sequence 7, Appli
24 32 38.6 445 9 US-10-076-622-473 Sequence 473, App
25 32 38.6 445 10 US-09-825-301-29 Sequence 29, Appli
26 32 38.6 445 10 US-09-604-287A-473 Sequence 473, App
27 32 38.6 445 12 US-10-007-805-473 Sequence 472, App
28 32 38.6 466 9 US-09-825-301-28 Sequence 472, App
29 32 38.6 466 10 US-09-604-287A-472 Sequence 472, App
30 32 38.6 466 12 US-10-007-805-472 Sequence 8, Appli
31 32 38.6 542 9 US-10-045-815-8 Sequence 50, Appli
32 32 38.6 542 10 US-09-159-469-50 Sequence 50, Appli
33 32 38.6 578 10 US-09-798-042-50 Sequence 469, App
34 32 38.6 650 9 US-10-076-622-469 Sequence 25, Appli
35 32 38.6 650 10 US-09-825-301-25 Sequence 25, Appli
36 32 38.6 650 10 US-09-604-287A-469 Sequence 469, App
37 32 38.6 650 12 US-10-007-805-469 Sequence 469, App
38 32 38.6 661 9 US-10-076-622-552 Sequence 552, App
39 32 38.6 661 12 US-10-007-805-552 Sequence 552, App
40 32 38.6 736 9 US-10-045-815-4 Sequence 4, Appli
41 32 38.6 743 9 US-10-076-622-494 Sequence 494, App
42 32 38.6 743 12 US-10-007-805-494 Sequence 494, App
43 32 38.6 747 9 US-10-045-815-6 Sequence 6, Appli
44 32 38.6 1002 9 US-10-076-622-475 Sequence 475, App
45 32 38.6

ALIGNMENTS

RESULT 1
US-09-925-301-1276
; Sequence 1276, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1276
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1276

Query Match 44.6%; Score 37; DB 10; Length 766;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 EEIVPNXVE 19
Db 595 EEVVPNVIE 603
||:||||:|

RESULT 2
US-09-864-761-37733
; Sequence 37733, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

```

1 FILE REFERENCE: Aeomica-x-1
2 CURRENT APPLICATION NUMBER: US/09/864,761
3 CURRENT FILING DATE: 2001-05-23
4 PRIOR APPLICATION NUMBER: US 60/180,312
5 PRIOR FILING DATE: 2000-02-04
6 PRIOR APPLICATION NUMBER: US 60/207,456
7 PRIOR FILING DATE: 2000-05-26
8 PRIOR APPLICATION NUMBER: US 09/632,366
9 PRIOR FILING DATE: 2000-08-03
10 PRIOR APPLICATION NUMBER: GB 24263.6
11 PRIOR FILING DATE: 2000-10-04
12 PRIOR APPLICATION NUMBER: US 60/236,359
13 PRIOR FILING DATE: 2000-09-27
14 PRIOR APPLICATION NUMBER: PCT/US01/006666
15 PRIOR FILING DATE: 2001-01-30
16 PRIOR APPLICATION NUMBER: PCT/US01/006667
17 PRIOR FILING DATE: 2001-01-30
18 PRIOR APPLICATION NUMBER: PCT/US01/006664
19 PRIOR FILING DATE: 2001-01-30
20 PRIOR APPLICATION NUMBER: PCT/US01/006669
21 PRIOR FILING DATE: 2001-01-30
22 PRIOR APPLICATION NUMBER: PCT/US01/006665
23 PRIOR FILING DATE: 2001-01-30
24 PRIOR APPLICATION NUMBER: PCT/US01/006668
25 PRIOR FILING DATE: 2001-01-30
26 PRIOR APPLICATION NUMBER: PCT/US01/006663
27 PRIOR FILING DATE: 2001-01-30
28 PRIOR APPLICATION NUMBER: PCT/US01/006662
29 PRIOR FILING DATE: 2001-01-30
30 PRIOR APPLICATION NUMBER: PCT/US01/006661
31 PRIOR FILING DATE: 2001-01-30
32 PRIOR APPLICATION NUMBER: PCT/US01/006670
33 PRIOR FILING DATE: 2001-01-30
34 PRIOR APPLICATION NUMBER: US 60/234,697
35 PRIOR FILING DATE: 2000-09-21
36 PRIOR APPLICATION NUMBER: US 09/608,408
37 PRIOR FILING DATE: 2000-06-30
38 PRIOR APPLICATION NUMBER: US 09/774,203
39 PRIOR FILING DATE: 2001-01-29
40 NUMBER OF SEQ ID NOS: 49117
41 SOFTWARE: Anomax Sequence Listing Engine
42 SEQ ID NO 37733
43 LENGTH: 253
44 TYPE: PRT
45 ORGANISM: Homo sapiens
46 FEATURE:
47 OTHER INFORMATION: MAP TO AC006059.3
48 OTHER INFORMATION: EXPRESSED IN HBL100, S
49 OTHER INFORMATION: EXPRESSED IN HELA, SIGH
50 OTHER INFORMATION: EXPRESSED IN PLACENTA,
51 OTHER INFORMATION: EXPRESSED IN HEART, SIGH
52 OTHER INFORMATION: EXPRESSED IN BT474, SIGH
53 OTHER INFORMATION: EXPRESSED IN FETAL LIV
54 OTHER INFORMATION: EXPRESSED IN BONE MARRO
55 OTHER INFORMATION: EXPRESSED IN LUNG, SIGH
56 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGH
57 OTHER INFORMATION: EST HUMAN HIT: A135094
58 OTHER INFORMATION: SWISSPROT HIT: P30414,
59 US-09-864-761-37733

```

```

: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA105
: CURRENT APPLICATION NUMBER: US/09/925,297
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05989
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 928
: SOFTWARE: PatentIn ver. 2.0
: SEQ ID NO 524
: LENGTH: 374
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (76)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (77)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (78)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: US-09-925-297-524

```

RESULT 3
US-09-925-297-524
; Sequence 524, Application US/09925297
; Patent No. US20020081659A1

; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Mather, Eric
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
; FILE REFERENCE: 09010/027001
; CURRENT APPLICATION NUMBER: US/09/391,340A
; CURRENT FILING DATE: 1999-09-07
; EARLIER APPLICATION NUMBER: US 08/907,166
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 803
; TYPE: PRT
; ORGANISM: Pyrolobus fumarius
; US-09-391-340-4

Query Match 41.0%; Score 34; DB 10; Length 803;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 EEIVPNXVEQ 20
:|||||:
Db 404 KELVPNRVER 413

RESULT 6
US-09-948-369-4
; Sequence 4, Application US/09948369
; Patent No. US20020132243A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: CALLEN, Walter
; APPLICANT: MATHUR, Eric
; TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DIVER1350-3
; CURRENT APPLICATION NUMBER: US/09/948,369
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US 09/656,309
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/391,340
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/907,166
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 803
; TYPE: PRT
; ORGANISM: Pyrolobus fumarius
; US-09-948-369-4

Query Match 41.0%; Score 34; DB 10; Length 803;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 EEIVPNXVEQ 20
:|||||:
Db 404 KELVPNRVER 413

RESULT 7
US-09-948-369-14
; Sequence 14, Application US/09948369
; Patent No. US20020132243A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: CALLEN, Walter
; APPLICANT: MATHUR, Eric
; TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DIVER1350-3

; CURRENT APPLICATION NUMBER: US/09/948,369
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US 09/656,309
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/391,340
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/907,166
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 803
; TYPE: PRT
; ORGANISM: Pyrolobus fumarius
; US-09-948-369-14

Query Match 41.0%; Score 34; DB 10; Length 803;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 EEIVPNXVEQ 20
:|||||:
Db 404 KELVPNRVER 413

RESULT 8
US-09-948-369-16
; Sequence 16, Application US/09948369
; Patent No. US20020132243A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: CALLEN, Walter
; APPLICANT: MATHUR, Eric
; TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY AND METHODS
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DIVER1350-3
; CURRENT APPLICATION NUMBER: US/09/948,369
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US 09/656,309
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/391,340
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: US 08/907,166
; PRIOR FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 803
; TYPE: PRT
; ORGANISM: Pyrolobus fumarius
; US-09-948-369-16

Query Match 41.0%; Score 34; DB 10; Length 803;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 EEIVPNXVEQ 20
:|||||:
Db 404 KELVPNRVER 413

RESULT 9
US-10-034-849-2
; Sequence 2, Application US/10034849
; Patent No. US20020115108A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: CALLEN, Walter
; APPLICANT: MATHUR, Eric
; TITLE OF INVENTION: ENZYMES HAVING HIGH TEMPERATURE POLYMERASE ACTIVITY
; TITLE OF INVENTION: AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: DIVER1350-2

CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078

Query Match 41.0%; Score 34; DB 10; Length 1217;
Best Local Similarity 38.9%;
Matches 7; Conservative 3; Mismatches 8; Indels
Pred. NO. 1.5e+02;

Matches	7;	Conservative	3;	Mismatches	8;	Indels	0;	Gaps	0;
Matches	7;	Conservative	3;	Mismatches	8;	Indels	0;	Gaps	0;

QY 3 EAEXIXXXEIVPNXVEQ 20
||| | | : ||
Db 134 EAENILLLEQTIDEVIEQ 151

RESULT 13

US-09-815-242-10892
; Sequence 10892, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10892
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10892

Query Match 39.8%; Score 33; DB 10; Length 305;
Best Local Similarity 36.8%; Pred. No. 45;
Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 QMEAXIXXXEIVPNXVE 19
| : | : | : ||| |
Db 172 QVPEELLNVTDIVPNETE 190

RESULT 14

US-08-834-666A-2
; Sequence 2, Application US/08834666A
; Patent No. US20020044949A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Lissolo, Ling
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Miller, Charles
; APPLICANT: Al-Garawi, Amal
; TITLE OF INVENTION: 76 kDa Helicobacter Polypeptides and
; TITLE OF INVENTION: Corresponding Polynucleotide Molecules
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA

Query Match 39.8%; Score 33; DB 10; Length 1781;
Best Local Similarity 38.9%; Pred. No. 3.6e+02;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,666A
; FILING DATE: 01-APR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,175
; REFERENCE/DOCKET NUMBER: 06132/038001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 708 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Signal Sequence
; LOCATION: 1...19
; OTHER INFORMATION:
US-08-834-666A-2

Query Match 39.8%; Score 33; DB 8; Length 708;
Best Local Similarity 37.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 QMEAXIXXXEIVPN 16
| : | : | : | : | : |
Db 354 QAQAEILNQAEQVVKV 369

RESULT 15

US-09-738-877-3
; Sequence 3, Application US/09738877
; Patent No. US20020015970A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan
; APPLICANT: Weiss, Stephen J.
; APPLICANT: Glynn, Richard
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF ANGIOGENESIS, COMPOSITIONS, AND
; TITLE OF INVENTION: SCREENING FOR ANGIOGENESIS MODULATORS
; FILE REFERENCE: A-69806/DJB/JJD
; CURRENT APPLICATION NUMBER: US/09/738,877
; CURRENT FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,425
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT/US 00/22061
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-738-877-3

Query Match 39.8%; Score 33; DB 10; Length 1781;
Best Local Similarity 38.9%; Pred. No. 3.6e+02;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEEIVPNXV 18
|:| | | | | | | | | |
Db 923 QVEAEALLTEEVLEREV 940

Search completed: February 11, 2003, 18:36:13
Job time : 12.1753 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:17:01 ; Search time 144.186 Seconds
(without alignments)
93.903 Million cell updates/sec

Title: US-09-380-738A-1

Perfect score: 83

Sequence: 1 QMEAEIXXXEIVPNXVEQK 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending_Patents_AA_Main.*
- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
 - 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
 - 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
 - 4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
 - 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
 - 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
 - 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
 - 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
 - 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
 - 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
 - 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
 - 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
 - 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
 - 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
 - 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
 - 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
 - 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
 - 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
 - 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
 - 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
 - 21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
 - 22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
 - 23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
 - 24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
 - 25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
 - 26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
 - 27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	94.0	125	9 US-08-545-573-13	Sequence 13, Appl
2	78	94.0	125	9 US-08-545-573A-13	Sequence 13, Appl
3	78	94.0	192	9 US-08-545-573-9	Sequence 9, Appl
4	78	94.0	192	9 US-08-545-573A-9	Sequence 9, Appl
5	78	94.0	199	9 US-08-545-573A-42	Sequence 42, Appl
6	78	94.0	199	14 US-09-066-330-9	Sequence 9, Appl

7	78	94.0	199	26	US-10-229-066-9	Sequence 9, Appl
8	78	94.0	214	9	US-08-545-573A-41	Sequence 41, Appl
9	75	90.4	21	3	US-07-731-592B-1	Sequence 1, Appl
10	73	88.0	21	17	US-09-380-738A-1	Sequence 1, Appl
11	63.5	76.5	124	9	US-08-545-573-15	Sequence 15, Appl
12	63.5	76.5	124	9	US-08-545-573A-15	Sequence 15, Appl
13	56.5	68.1	200	9	US-08-545-573-21	Sequence 21, Appl
14	56.5	68.1	200	9	US-08-545-573-23	Sequence 23, Appl
15	56.5	68.1	200	9	US-08-545-573A-21	Sequence 21, Appl
16	56.5	68.1	200	9	US-08-545-573A-23	Sequence 23, Appl
17	39	47.0	1280	21	US-09-708-427-11897	Sequence 11897, A
18	39	47.0	1280	23	US-09-935-625-16364	Sequence 16364, A
19	39	47.0	1280	23	US-09-935-625-28072	Sequence 28072, A
20	39	47.0	1286	23	US-09-935-625-18244	Sequence 18244, A
21	39	47.0	1286	23	US-09-935-625-30821	Sequence 30821, A
22	39	47.0	1328	21	US-09-708-427-11896	Sequence 11896, A
23	39	47.0	1328	23	US-09-935-625-16363	Sequence 16363, A
24	39	47.0	1328	23	US-09-935-625-28071	Sequence 28071, A
25	39	47.0	1334	23	US-09-935-625-18243	Sequence 18243, A
26	39	47.0	1334	23	US-09-935-625-30820	Sequence 30820, A
27	39	47.0	1396	21	US-09-708-427-11895	Sequence 11895, A
28	39	47.0	1396	23	US-09-935-625-16362	Sequence 16362, A
29	39	47.0	1396	23	US-09-935-625-28070	Sequence 28070, A
30	39	47.0	1402	23	US-09-935-625-18242	Sequence 18242, A
31	39	47.0	1402	23	US-09-935-625-30819	Sequence 30819, A
32	38	45.8	14	16	US-09-222-791-1	Sequence 1, Appl
33	38	45.8	131	16	US-09-248-796-23250	Sequence 23250, A
34	38	45.8	131	27	US-60-096-409-23250	Sequence 23250, A
35	38	45.8	240	21	US-09-791-537-105233	Sequence 105233, A
36	38	45.8	466	20	US-09-614-150-2832	Sequence 2832, Ap
37	38	45.8	466	27	US-60-167-217-2880	Sequence 2880, Ap
38	38	45.8	466	27	US-60-191-637-2838	Sequence 2838, Ap
39	38	45.8	535	19	US-09-570-581A-1366	Sequence 1366, Ap
40	38	45.8	2421	20	US-09-614-150-23721	Sequence 23721, A
41	38	45.8	2421	27	US-60-191-637-23807	Sequence 23807, A
42	38	45.8	2421	27	US-60-191-681-18758	Sequence 18758, A
43	38	45.8	2887	27	US-60-173-484-19523	Sequence 19523, A
44	37	44.6	392	21	US-09-791-537-3729	Sequence 3729, Ap
45	37	44.6	530	22	US-09-897-516-6110	Sequence 6110, Ap

ALIGNMENTS

RESULT 1

US-08-545-573-13
Sequence 13, Application US/08545573
GENERAL INFORMATION:
APPLICANT: CARR, N.G.
TITLE OF INVENTION: PHENYLALANINE-FREE PROTEIN AND DNA CODING
TITLE OF INVENTION: THEREFORE
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/545,573
FILING DATE: 16-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO GB94/01046
FILING DATE: 16-MAY-1994
PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: GB 9310472.7
; FILING DATE: 20-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LIEBESCHUETZ, JOE
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 016994-0119000US
; TELEPHONE: (650) 326-2400
; TELEFAX: (650) 326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 125 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-545-573A-13

Query Match          94.0%; Score 78; DB 9; Length 125;
Best Local Similarity 76.2%; Pred. No. 2e-07;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEXIXXXEIVPNXVEQK 21
   ||||| | ||||| |||||
Db 66 QMEAESISSSEIVPNSVEQK 86

RESULT 2
US-08-545-573A-13
; Sequence 13, Application US/08545573A
; GENERAL INFORMATION:
; APPLICANT: Carr, Noel Gordon
; TITLE OF INVENTION: Phenylalanine-Free Protein and DNA Coding
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,573A
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/01046
; FILING DATE: 16-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9310472.7
; FILING DATE: 20-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 016994-0119000US
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 125 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..125
```

```
; OTHER INFORMATION: /note= "predicted sequence of bovine
; OTHER INFORMATION: casein block A"
; US-08-545-573A-13

Query Match          94.0%; Score 78; DB 9; Length 125;
Best Local Similarity 76.2%; Pred. No. 2e-07;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEXIXXXEIVPNXVEQK 21
   ||||| | ||||| |||||
Db 66 QMEAESISSSEIVPNSVEQK 86

RESULT 3
US-08-545-573-9
; Sequence 9, Application US/08545573
; GENERAL INFORMATION:
; APPLICANT: Carr, N.G.
; APPLICANT: MANN, N.H.
; TITLE OF INVENTION: PHENYLALANINE-FREE PROTEIN AND DNA CODING
; TITLE OF INVENTION: THEREFORE
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,573
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO GB94/01046
; FILING DATE: 16-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9310472.7
; FILING DATE: 20-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LIEBESCHUETZ, JOE
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 016994-0119000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 326-2400
; TELEFAX: (650) 326-2422
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 192 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..192
; OTHER INFORMATION: /label= protein
; OTHER INFORMATION: /note= "mature bovine alpha-s1-casein with
; OTHER INFORMATION: phenylalanine residues removed and addition of
; OTHER INFORMATION: methionine residue."
; US-08-545-573-9

Query Match          94.0%; Score 78; DB 9; Length 192;
Best Local Similarity 76.2%; Pred. No. 3.6e-07;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEXIXXXEIVPNXVEQK 21
   ||||| | ||||| |||||
Db 56 QMEAESISSSEIVPNSVEQK 76
```



```

RESULTS 4
US-08-545-573A-9
; Sequence 9, Application US/08545573A
; GENERAL INFORMATION:
; APPLICANT: Carr, Noel Gordon
; APPLICANT: Mann, Nicholas Harold
; TITLE OF INVENTION: Phenylalanine-Free Protein and DNA Coding
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545.573A
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/01046
; FILING DATE: 16-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9310472.7
; FILING DATE: 20-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 016994-0119000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 192 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..192
; OTHER INFORMATION: /note= "mature bovine alpha-s1-casein
; OTHER INFORMATION: with phenylalanine residues removed and
; OTHER INFORMATION: addition of an N-terminal methionine
; OTHER INFORMATION: residue"
US-08-545-573A-9
; Query Match 94.0%; Score 78; DB 9; Length 192;
; Best Local Similarity 76.2%; Pred. No. 3.6e-07;
; Matches 16; Conservative 0; Mismatches 5; Indels 0;

QY 1 QMEAEIXXXEIVPNKVEQK 21
||||| | ||||||| |||||

DB 56 QMEAESISSSEIIVPNSVEQK 76

RESULTS 5
US-08-545-573A-42
; Sequence 42, Application US/08545573A
; GENERAL INFORMATION:
; APPLICANT: Carr, Noel Gordon
; APPLICANT: Mann, Nicholas Harold
; TITLE OF INVENTION: Phenylalanine-Free Protein and DNA Coding
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 43

```

Query Match 94.0%; Score 78; DB 14; Length 199;
Best Local Similarity 76.2%; Pred. No. 3.8e-07;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
||||| | ||||| |||||
DB 59 QMEAESISSSEIVPNXVEQK 79

RESULT 7

US-10-229-066-9
; Sequence 9, Application US/10229066
; GENERAL INFORMATION:
; APPLICANT: Reynolds, Eric C.
; APPLICANT: Bhogal, Peter S.
; APPLICANT: Slakeski, Nada
; TITLE OF INVENTION: DIAGNOSTICS AND TREATMENTS OF PERIODONTAL DISEASE
; FILE REFERENCE: Reynolds
; CURRENT APPLICATION NUMBER: US/10/229,066
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US/09/066,330
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: PN 6275
; PRIOR FILING DATE: 1995-10-30
; PRIOR APPLICATION NUMBER: PCT/AU96/00673
; PRIOR FILING DATE: 1996-10-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-229-066-9

Query Match 94.0%; Score 78; DB 26; Length 199;
Best Local Similarity 76.2%; Pred. No. 3.8e-07;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
||||| | ||||| |||||
DB 59 QMEAESISSSEIVPNXVEQK 79

RESULT 8

US-08-545-573A-41
; Sequence 41, Application US/08545573A
; GENERAL INFORMATION:
; APPLICANT: Carr, Noel Gordon
; APPLICANT: Mann, Nicholas Harold
; TITLE OF INVENTION: Phenylalanine-Free Protein and DNA Coding
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,573A
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/01046
; FILING DATE: 16-MAY-1994
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9310472.7
; FILING DATE: 20-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 016994-0119000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-545-573A-41

Query Match 94.0%; Score 78; DB 9; Length 214;
Best Local Similarity 76.2%; Pred. No. 4.2e-07;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
||||| | ||||| |||||
DB 74 QMEAESISSSEIVPNXVEQK 94

RESULT 9

US-07-731-592B-1
; Sequence 1, Application US/07731592B
; GENERAL INFORMATION:
; APPLICANT: Burger, Allan R.
; APPLICANT: Elliott, David L.
; APPLICANT: Schick, Laura A.
; TITLE OF INVENTION: Oral Compositions Containing a
; TITLE OF INVENTION: Phosphopeptide
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Department, Unilever United States, Inc.
; STREET: 45 River Road
; CITY: Edgewater
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07020

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette-3.5 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Microsoft Word(R)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/731,592B
; FILING DATE: 19910717
; CLASSIFICATION: 514
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: AMINO ACID
; FEATURE:

; NAME/KEY: Phosphoserine
; LOCATION: 6
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 8
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 9
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
; FEATURE:
; NAME/KEY: Phosphoserine

us-09-380-738a-1.rapm

wed Feb 12 11:35:20 2003

```

; LOCATION: 10
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 17
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
; US-07-731-592B-1
;
; Query Match 90.4%; Score 75; DB 3; Length 21;
; Best Local Similarity 71.4%; Pred. No. 7.4e-08;
; Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
;
; QY 1 OMEAEIXXXEIVPNXVEQK 21
; :||||| : ||||| |||||
; Db 1 EMEAESISSEIIVPNSVEQK 21
;
; RESULT 10
; US-09-380-738A-1
; Sequence 1, Application US/09380738A
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, Eric
; TITLE OF INVENTION: CALCIUM PHOSHOPEPTIDE COMPLEXES
; FILE REFERENCE: 040268/0161
; CURRENT APPLICATION NUMBER: US/09/380,738A
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: PCT/AU98/00160
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: AU P05662
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Bos sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (6)..(6)
; OTHER INFORMATION: Xaa is a phosphorylated serine
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: Xaa is a phosphorylated serine
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Xaa is a phosphorylated serine
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Xaa is a phosphorylated serine
; NAME/KEY: misc_feature
; LOCATION: (17)..(17)
; OTHER INFORMATION: Xaa is a phosphorylated serine
; US-09-380-738A-1
;
; Query Match 88.0%; Score 73; DB 17; Length 21;
; Best Local Similarity 100.0%; Pred. No. 1.9e-07;
; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 OMEAEIXXXEIVPNXVEQK 21
; :||||| : ||||| ||||| |||||
; Db 1 OMEAEIXXXEIVPNXVEQK 21
;
; RESULT 11
; US-08-545-573-15
; Sequence 15, Application US/08545573
; GENERAL INFORMATION:
; APPLICANT: CARR, N.G.
; TITLE OF INVENTION: PHENYLALANINE-FREE PROTEIN AND DNA CODING
; THEREFORE
;
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,573
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO GB94/01046
; FILING DATE: 16-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9310472.7
; FILING DATE: 20-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LIEBSCHUETZ, JOE
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 016994-0119000S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 326-2400
; TELEFAX: (650) 326-2422
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-545-573-15
;
; Query Match 76.5%; Score 63.5; DB 9; Length 124;
; Best Local Similarity 71.4%; Pred. No. 0.00019;
; Matches 15; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
;
; QY 1 OMEAEIXXXEIVPNXVEQK 21
; :||||| : ||||| |||||
; Db 66 OMEAESISSEIIVPNS-EQK 85
;
; RESULT 12
; US-08-545-573A-15
; Sequence 15, Application US/08545573A
; GENERAL INFORMATION:
; APPLICANT: Carr, Noel Gordon
; APPLICANT: Mann, Nicholas Harold
; TITLE OF INVENTION: Phenylalanine-Free Protein and DNA Coding
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,573A
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 435

```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/01046
; FILING DATE: 16-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9310472.7
; FILING DATE: 20-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 016994-01190005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..124
; OTHER INFORMATION: /note= "actual sequence obtained for
; OTHER INFORMATION: bovine casein block A"
US-08-545-573A-15

Query Match          76.5%; Score 63.5; DB 9; Length 124;
Best Local Similarity 71.4%; Pred. No. 0.00019;
Matches 15; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 QMEAEIXXXEIVPNXVEQK 21
Db 66 QMEAESISSSEIVPNS-EQK 85

RESULT 13
US-08-545-573-21
; Sequence 21, Application US/08545573
; GENERAL INFORMATION:
; APPLICANT: CARR, N.G.
; TITLE OF INVENTION: PHENYLALANINE-FREE PROTEIN AND DNA CODING
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,573
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO GB94/01046
; FILING DATE: 16-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9310472.7
; FILING DATE: 20-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LIEBESCHUETZ, JOE
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 016994-01190005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 326-2400
; TELEFAX: (650) 326-2422
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 200 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-545-573-23

Query Match          68.1%; Score 56.5; DB 9; Length 200;
Best Local Similarity 66.7%; Pred. No. 0.0097;
Matches 14; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
```

```
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 200 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-545-573-21

Query Match          68.1%; Score 56.5; DB 9; Length 200;
Best Local Similarity 66.7%; Pred. No. 0.0097;
Matches 14; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 QMEAEIXXXEIVPNXVEQK 21
Db 66 QMEAESISSSEIVPNS-EQK 85

RESULT 14
US-08-545-573-23
; Sequence 23, Application US/08545573
; GENERAL INFORMATION:
; APPLICANT: CARR, N.H.
; TITLE OF INVENTION: PHENYLALANINE-FREE PROTEIN AND DNA CODING
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,573
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO GB94/01046
; FILING DATE: 16-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9310472.7
; FILING DATE: 20-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LIEBESCHUETZ, JOE
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 016994-01190005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 326-2400
; TELEFAX: (650) 326-2422
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 200 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-545-573-23

Query Match          68.1%; Score 56.5; DB 9; Length 200;
Best Local Similarity 66.7%; Pred. No. 0.0097;
Matches 14; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
```

us-09-380-738a-1.1.rapm

Wed Feb-12 11:35:20 2003

```

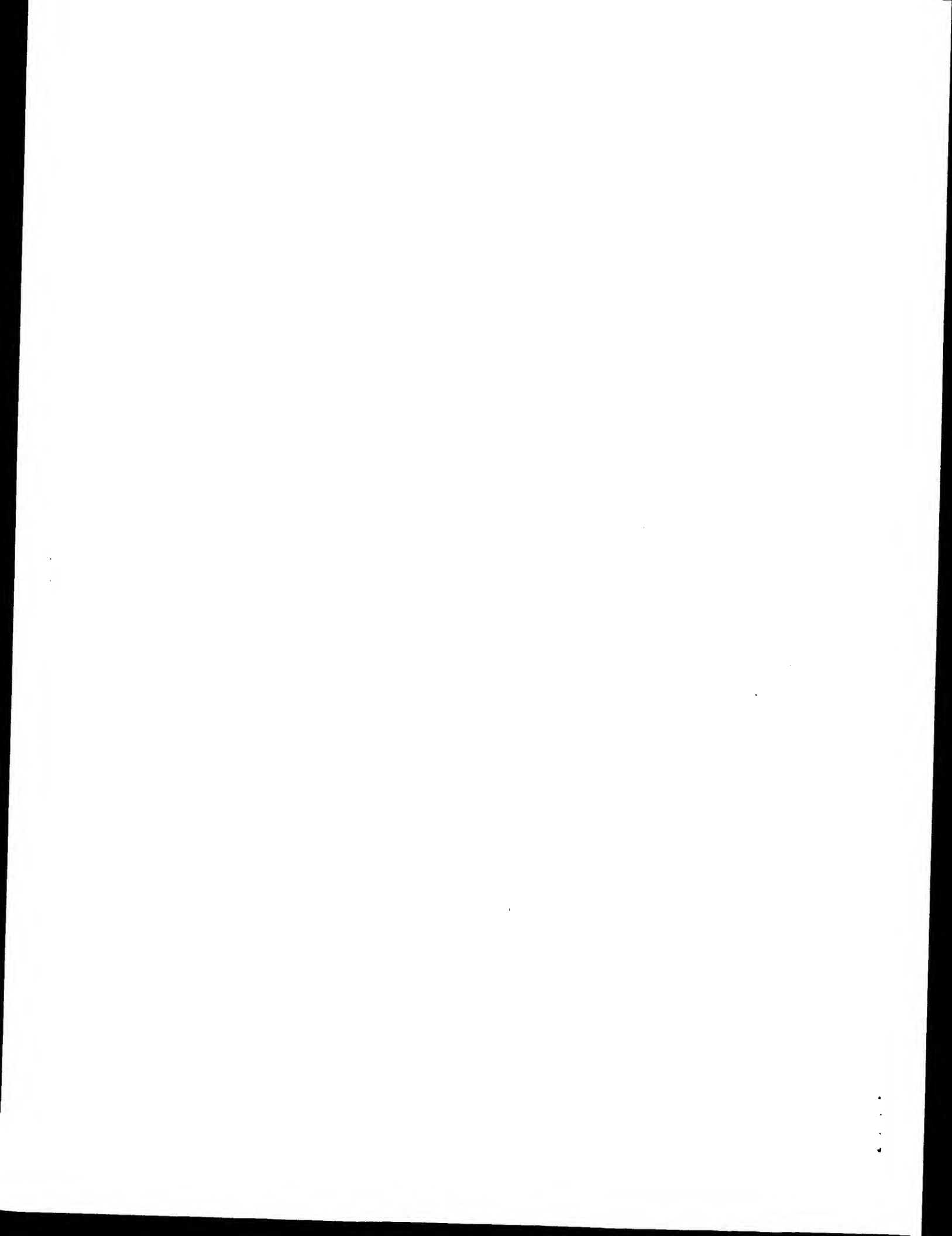
RESULT 15
US-08-545-573A-21
; Sequence 21 Application US/08545573A
; GENERAL INFORMATION:
; APPLICANT: Carr, Noel Gordon
; TITLE OF INVENTION: Phenylalanine-Free Protein and DNA Coding
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,573A
; FILING DATE: 16-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB94/01046
; FILING DATE: 16-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9310472.7
; FILING DATE: 20-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 016994-0119000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 200 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..200
; OTHER INFORMATION: /note= "combined sequence of bovine
; OTHER INFORMATION: caseln blocks A and B"
US-08-545-573A-21

Query Match 68.1%; Score 56.5; DB 9; Length 200;
Best Local Similarity 66.7%; Pred. NO. 0.0097;
Matches 14; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 QMEAEIXXXEIVPNXVEOK 21
Db 66 QNEAESISSSEIVPNS-EOK 85

Search completed: February 11, 2003, 18:33:20
Job time : 146.186 secs

```



Wed Feb 12 11:35:21 2003

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:17:56 ; Search time 24.0309 seconds
(without alignments)
72.734 Million cell updates/sec

Title: US-09-380-738A-1
Perfect score: 83
Sequence: 1 QMEAEIXXXEIVPNKVEQK 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 420401 seqs, 83231269 residues

Total number of hits satisfying chosen parameters: 420401

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/pdata/2/paa/pct_NEW_COMB.pep.*
2: /cgn2_6/pdata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/pdata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/pdata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/pdata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/pdata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/pdata/2/paa/US60_NEW_COMB.pep.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	43.4	96	5	US-09-724-676-82488
2	36	43.4	96	5	US-09-724-676-82488
3	36	43.4	117	5	US-09-724-676-82494
4	36	43.4	117	5	US-09-724-676-82495
5	36	43.4	117	5	US-09-724-676-82494
6	36	43.4	117	5	US-09-724-676-82495
7	36	43.4	161	5	US-09-724-676-82492
8	36	43.4	161	5	US-09-724-676-82493
9	36	43.4	161	5	US-09-724-676-82492
10	36	43.4	161	5	US-09-724-676-82493
11	36	43.4	1265	5	US-09-724-676-82498
12	36	43.4	1265	5	US-09-724-676-82498
13	36	43.4	1274	5	US-09-724-676-82477
14	36	43.4	1274	5	US-09-724-676-82477
15	36	43.4	1280	5	US-09-724-676-82483
16	36	43.4	1280	5	US-09-724-676-82489
17	36	43.4	1302	5	US-09-724-676-82489
18	36	43.4	1302	5	US-09-724-676-82496
19	36	43.4	1336	5	US-09-724-676-82497
20	36	43.4	1336	5	US-09-724-676-82496
21	36	43.4	1336	5	US-09-724-676A-82497
22	36	43.4	1336	1	PCT-US02-04915-215
23	36	43.4	1338	6	US-10-293-017-56
24	36	43.4	1338	6	US-10-293-071-56
25	36	43.4	1339	6	US-10-165-193A-9
26	36	43.4	1339	6	US-10-165-193A-9

Sequence 82475, A
Sequence 82476, A
Sequence 82475, A
Sequence 82476, A
Sequence 82481, A
Sequence 82482, A
Sequence 82481, A
Sequence 82482, A
Sequence 82482, A
Sequence 82486, A
Sequence 82487, A
Sequence 82486, A
Sequence 82487, A
Sequence 14852, A
Sequence 381, App
Sequence 17, Appl
Sequence 495, App
Sequence 2, Appl
Sequence 8, Appl
Sequence 146, App

ALIGNMENTS

RESULT 1
US-09-724-676-82488
; Sequence 82488, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82488
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-82488

Query Match 43.4%; Score 36; DB 5; Length 96;
Best Local Similarity 46.7%; Pred. No. 4.3;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 MEAEIXXXEIVPN 16
Db 43 MSLEIKTEBLLPN 57

RESULT 2
US-09-724-676A-82488
; Sequence 82488, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82488
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-82488

Query Match 43.4%; Score 36; DB 5; Length 96;
Best Local Similarity 46.7%; Pred. No. 4.3;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 MEAEIXXXEIVPN 16

wed Feb 12 11:35:21 2003

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-82493

Query Match      43.4%; Score 36; DB 5; Length 161;
Best Local Similarity 46.7%; Pred. No. 8;
Matches 7; Conservative 2; Mismatches 6; Indels 6; Gaps 0;

QY 2 MEAEXIXXXEIVPN 16
Db 43 MSLEIKTFEELLPN 57

RESULT 9
US-09-724-676A-82492
; Sequence 82492, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82492
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-82492

Query Match      43.4%; Score 36; DB 5; Length 161;
Best Local Similarity 46.7%; Pred. No. 8;
Matches 7; Conservative 2; Mismatches 6; Indels 6; Gaps 0;

QY 2 MEAEXIXXXEIVPN 16
Db 43 MSLEIKTFEELLPN 57

RESULT 10
US-09-724-676A-82493
; Sequence 82493, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82493
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-82493

Query Match      43.4%; Score 36; DB 5; Length 161;
Best Local Similarity 46.7%; Pred. No. 8;
Matches 7; Conservative 2; Mismatches 6; Indels 6; Gaps 0;

QY 2 MEAEXIXXXEIVPN 16
Db 43 MSLEIKTFEELLPN 57

RESULT 11
US-09-724-676-82498
; Sequence 82498, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82498
; LENGTH: 1265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-82498

Query Match      43.4%; Score 36; DB 5; Length 1265;
Best Local Similarity 46.7%; Pred. No. 98;
Matches 7; Conservative 2; Mismatches 6; Indels 6; Gaps 0;

QY 2 MEAEXIXXXEIVPN 16
Db 1220 MSLEIKTFEELLPN 1234

RESULT 12
US-09-724-676A-82498
; Sequence 82498, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82498
; LENGTH: 1265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-82498

Query Match      43.4%; Score 36; DB 5; Length 1265;
Best Local Similarity 46.7%; Pred. No. 98;
Matches 7; Conservative 2; Mismatches 6; Indels 6; Gaps 0;

QY 2 MEAEXIXXXEIVPN 16
Db 1220 MSLEIKTFEELLPN 1234

RESULT 13
US-09-724-676-82477
; Sequence 82477, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82477
; LENGTH: 1274
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-82477

Query Match      43.4%; Score 36; DB 5; Length 1274;
Best Local Similarity 46.7%; Pred. No. 99;
Matches 7; Conservative 2; Mismatches 6; Indels 6; Gaps 0;

QY 2 MEAEXIXXXEIVPN 16
Db 1242 MSLEIKTFEELLPN 1256

RESULT 14
US-09-724-676A-82477
; Sequence 82477, Application US/09724676A
```

```
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82477
; LENGTH: 1274
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-82477
```

```
Query Match 43.4%; Score 36; DB 5; Length 1274;
Best Local Similarity 46.7%; Pred. No. 99;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Caps 0;
```

```
QY 2 MEAEXIXXXEIVPN 16
| | | | | | | | | | | | | | | | | | | | | |
Db 1242 MSLEIKTFEELLPN 1256
```

RESULT 15

```
US-09-724-676-82483
; Sequence 82483, Application US/09724676
; GENERAL INFORMATION:
```

```
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82483
; LENGTH: 1280
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-82483
```

```
Query Match 43.4%; Score 36; DB 5; Length 1280;
Best Local Similarity 46.7%; Pred. No. 99;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 2 MEAEXIXXXEIVPN 16
| | | | | | | | | | | | | | | | | | | | | |
Db 1248 MSLEIKTFEELLPN 1262
```

Search completed: February 11, 2003, 18:35:17
Job time : 25.0309 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: February 11, 2003, 18:14:11 ; Search time 18.4021 seconds
(without alignments)
109.706 Million cell updates/sec

Title: US-09-380-738A-1
Perfect score: 83
Sequence: 1 QMEAEIXXXEIVPNXBQK 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	94.0	214	1 KABOSB	alpha-s1-casein pr
2	57	68.7	206	1 KASHS1	alpha-s1-casein pr
3	57	68.7	214	2 S34262	alpha-s1-casein pr
4	55	66.3	198	2 S07130	casein B precursor
5	54	65.1	214	2 S06102	alpha-s1-casein -
6	44	53.0	215	2 JC1133	alpha-s1-casein pr
7	44	53.0	814	2 T23200	hypothetical prote
8	41	49.4	93	2 T13384	FMFamide-like pep
9	39	47.0	313	2 A36060	alpha-casein prote
10	39	47.0	323	2 C70205	hypothetical prote
11	39	47.0	508	2 T20355	hypothetical prote
12	39	47.0	1396	2 T10627	hypothetical prote
13	38	45.8	113	2 S66805	hypothetical prote
14	38	45.8	388	2 JC5164	circumsporozoite p
15	38	45.8	466	2 A37351	IMP-E2 protein pre
16	37	44.6	157	2 F72612	hydrolyase, ama/hip
17	37	44.6	357	2 F72245	glycine-tRNA ligas
18	37	44.6	739	2 A55314	DNA polymerase I l
19	36.5	44.0	875	2 AE1270	DNA polymerase I l
20	36.5	44.0	875	2 AG1632	alpha s1-casein -
21	36	43.4	206	1 A45661	hypothetical prote
22	36	43.4	234	2 D95021	hypothetical prote
23	36	43.4	234	2 A97893	transcription regu
24	36	43.4	466	2 A11707	mismatch repair pr
25	36	43.4	840	2 B86901	hypothetical prote
26	36	43.4	1075	2 T38253	protein-tyrosine k
27	36	43.4	1338	2 S09882	genome polyprotein
28	35.5	42.8	3396	1 A42551	conserved hypothet
29	35	42.2	129	2 C95164	

ALIGNMENTS

RESULT 1

KABOSB

alpha-s1-casein precursor - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 24-Apr-1984 #sequence_revision 06-Feb-1995 #text_change 21-Jul-2000

C:Accession: S22575; A23071; S02202; A90940; A91188; A91274; S23156; A90738;

R:Koczan, D.; Hobom, G.; Seyfert, H.M.

Nucleic Acids Res. 19, 5591-5596, 1991

A:Title: Genomic organization of the bovine alpha-s1 casein gene.

A:Reference number: S22575; MUID:92051301; PMID:1658736

A:Accession: S22575

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-214 <KOC>

A:Cross-references: EMBL:X59856; NID:g91; PIDN:CAA42516.1; PID:g92

R:Stewart, A.F.; Willis, I.M.; Mackinlay, A.G.

Nucleic Acids Res. 12, 3895-3907, 1984

A:Title: Nucleotide sequences of bovine alpha-s1- and kappa-casein cDNAs.

A:Reference number: A93517; MUID:84221403; PMID:6328443

A:Accession: A23071

A:Molecule type: mRNA

A:Residues: 1-214 <STE>

A:Cross-references: GB:X00564; NID:g175; PIDN:CAB57792.1; PID:g6015490

R:Gorodetskiy, S.I.; Zakhariev, V.M.; Kyarshulite, D.R.; Kapelinskaya, T.V.; Skryab.

Biochemistry (N.Y.) 51, 1402-1409, 1986

A:Title: Cloning and nucleotide sequence of cDNA for bovine alpha(S1)-casein.

A:Reference number: S02202

A:Accession: S02202

A:Molecule type: mRNA

A:Residues: 1-142, 'D', 144-210, 'IS', 213-214 <GOR>

A:Cross-references: EMBL:M38641

A:Note: this paper is a translation of the Russian paper published in Biokhimiya (1986)

A:Note: the authors translated the codon CAA for residue 74 as Glu, GAG for residue

R:Mercier, J.C.; Grosclaude, F.; Ribadeau-Dumas, B.

Eur. J. Biochem. 40, 323, 1973

A:Title: Structure primaire de la caseine alpha-s1 et de la caseine beta bovine.

A:Reference number: A91214; MUID:74082545; PMID:4797901

A:Contents: annotation; A, B, C, and D variants, revisions to positions 74 and 92-9

R:Willis, I.M.; Stewart, A.F.; Caputo, A.; Thompson, A.R.; Mackinlay, A.G.

DNA 1, 375-386, 1982

A:Title: Construction and identification by partial nucleotide sequence analysis of

A:Reference number: A90940; MUID:83182023; PMID:6897774

A:Accession: A90940

A:Molecule type: mRNA

A:Residues: 55-130 <WIL>

A:Cross-references: GB:K01084

R:Mercier, J.C.; Grosclaude, F.; Ribadeau-Dumas, B.

Eur. J. Biochem. 23, 41-51, 1971

A:Title: Structure primaire de la caseine alpha-s1-bovine. Sequence complete.

A:Reference number: A91188; MUID:72063417; PMID:4331376

A:Contents: B variant

A:Accession: A91188

conserved hypothet
virB6 protein - Ag
component of type
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
oxidative cyclase
CT651 hypothetical
CT651 hypothetical
conserved hypothet
transketolase [EC
transketolase [EC
hypothetical prote

A:Molecule type: protein
A:Residues: 16-44,'Q',46-214 <MER>
A:Note: the sequences of residues 85-99 and of 125-140 are similar, having 10 identities
R:Grosclaude, F.; Mahe, M.F.; Mercier, J.C.; Ribadeau-Dumas, B.
Eur. J. Biochem. 26, 328-337, 1972
A:Title: Caractérisation des variants génétiques des caséines alpha-S1 et beta bovines.
A:Reference number: A91192; MUID:72214259; PMID:5064450
A:Contents: annotation; D variant
A:Accession: A91192
A:Molecule type: protein
A:Residues: 16-44,'Q',46-67,'T',69-214 <GRO>
A:Note: the D variant has a phosphorylated Thr at position 68
R:Grosclaude, F.; Mahe, M.F.; Mercier, J.C.; Ribadeau-Dumas, B.
FEBS Lett. 11, 109-112, 1970
A:Title: Localisation, dans la partie NH-2-terminale de la caséine alpha-S1 bovine, d'un
A:Reference number: A91274
A:Contents: A variant
A:Accession: A91274
A:Molecule type: protein
A:Residues: 23-28,42-44,'Q',46-49 <GR2>
A:Note: the sequence of the A variant lacks residues 29-41 of the B variant
R:Neuteboom, B.; Gluifrida, M.G.; Conti, A.
FEBS Lett. 305, 189-191, 1992
A:Title: Isolation of a new ligand-carrying casein fragment from bovine mammary gland mi
A:Reference number: S23156; MUID:93231344; PMID:1299613
A:Accession: S23156
A:Molecule type: protein
A:Residues: 39-55 <NEU>
R:Grosclaude, F.; Mercier, J.C.; Ribadeau-Dumas, B.
C. R. Acad. Sci. Hebd. Seances Acad. Sci. D 268, 3133-3136, 1969
A:Title: Sur la localisation, dans la séquence COOH-terminale de la caséine alpha-S1 bo
A:Reference number: A90738; MUID:69257961; PMID:4979278
A:Contents: C variant
A:Accession: A90738
A:Molecule type: protein
A:Residues: 205-206,'G',208-214 <GR3>
R:Grosclaude, F.; Mercier, J.C.; Ribadeau-Dumas, B.
C. R. Acad. Sci. D Sci. Nat. 271, 563, 1970
A:Reference number: A90740
A:Contents: annotation; erratum: C variant, revision
R:Exterkate, F.A.; Altling, A.C.; Slangen, C.J.
Biochem. J. 273, 135-139, 1991
A:Title: Specificity of two genetically related cell-envelope proteinases of Lactococcus
A:Reference number: S13300; MUID:9113130; PMID:1899185
A:Contents: annotation; cleavage by Lactococcus lactis subsp. cremoris cell-envelope pro
R:Nagao, M.; Maki, M.; Sasaki, R.; Chiba, R.
Agric. Biol. Chem. 48, 1663-1667, 1984
A:Title: Isolation and sequence analysis of bovine alpha-S1-casein cDNA clone.
A:Reference number: I45871
A:Accession: I45871
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-214 <NAG>
A:Cross-references: GB:M33123; NID:g162791; PIDN:AAA30428.1; PID:g162792
R:Kiarshulite, D.R.; Zakhar'ev, V.M.; Gorodetskii, S.I.
Dokl. Akad. Nauk SSSR 280, 1433-1437, 1985
A:Title: [Nucleotide sequence of the 3'-nontranslated region of the mRNA of alpha S1-cas
A:Reference number: I45852; MUID:85178933; PMID:3838718
A:Accession: I45852
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 122-202,'L',204-214 <KIA>
A:Cross-references: GB:M38658; NID:g162649; PIDN:AAA62707.1; PID:g162650
R:Maki, M.; Nagao, M.; Hirose, M.; Chiba, H.
Agric. Biol. Chem. 47, 441-444, 1983
A:Title: Cloning of cDNA sequence coding for bovine alpha-S1-casein.
A:Reference number: I45870
A:Accession: I45870
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 185-202 <NAK>
A:Cross-references: GB:D00412; NID:g217532; PIDN:BAA00313.1; PID:g217533
R:Gorodetskii, S.I.; Zakhar'ev, V.M.; Kiarshulite, D.R.; Kapelinskaya, T.V.; Skryabin, N

Biochimia 51, 1402-1409, 1986
A:Title: Cloning and nucleotide sequence of cDNA for bovine alpha-S1-casein.
A:Reference number: I45872
A:Accession: I45872
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-41,'L',43-142,'D',144-210,'IS',213-214 <GO2>
A:Cross-references: GB:M38641; NID:g162793; PIDN:AAA30429.1; PID:g162794
C:Comment: The B variant sequence is shown.
C:Genetics:
A:Map position: 6
A:Introns: 17/3; 28/3; 41/3; 49/3; 57/3; 65/3; 73/3; 84/3; 92/3; 110/3; 124/3; 132/3; 132/3;
C:Superfamily: alpha-s1-casein
C:Keywords: mammary gland; milk; phosphoprotein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-214/Product: alpha-s1-casein #status predicted <MAT>
F:61,63,79,81,82,83,90,130/Binding site: phosphate (Ser) (covalent) #status experim
Query Match 94.0%; Score 78; DB 1; Length 214;
Best Local Similarity 76.2%; Pred. No. 1.5e-07;
Matches 16; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 QMEAEIXXXEEIVPNXVEQK 21
||||| | ||||| |||||
Db 74 QMEAESISSSEIEIVPNSVEQK 94
RESULT 2
KASHSI
alpha-s1-casein precursor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 22-Jun-1999
C:Accession: A25069
R:Mercier, J.C.; Gaye, P.; Soulier, S.; Hue-Delahaie, D.; Vilotte, J.L.
Biochimie 67, 959-971, 1985
A:Title: Construction and identification of recombinant plasmids carrying cDNAs codin
in cDNA.
A:Reference number: A25069; MUID:86104473; PMID:3002501
A:Accession: A25069
A:Molecule type: mRNA
A:Residues: 1-206 <MER>
A:Cross-references: GB:X03237; NID:g1247; PIDN:CAA26982.1; PID:g1248
A:Note: the authors translated the codon TCT for residue 27 as Asp
C:Genetics:
A:Map position: 2
C:Superfamily: alpha-s1-casein
C:Keywords: mammary gland; milk; phosphoprotein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-206/Product: alpha-s1-casein #status predicted <KAL>
F:105-111/Region: oploid-like peptide sequence
F:79,80,81,82,83/Binding site: phosphate (Ser) (covalent) #status predicted
Query Match 68.7%; Score 57; DB 1; Length 206;
Best Local Similarity 57.1%; Pred. No. 0.0018;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
QY 1 QMEAEIXXXEEIVPNXVEQK 21
||||| | ||||| |||||
Db 74 QMKAGSSSSSEIEIVPNSAEQK 94
RESULT 3
S34262
alpha-s1-casein - goat
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
C:Accession: S34262
R:Jansa, M.; Leroux, C.; Sanchez, A.; Martin, P.
submitted to the EMBL Data Library, May 1993
A:Description: Occurrence of a LINE element in the 3' UTR of an allelic form of the go
A:Reference number: S34262
A:Accession: S34262
A>Status: preliminary

Wed Feb 12 11:35:21 2003

us-09-380-738a-1.rpr

A:Molecule type: mRNA

A:Residues: 1-214 <JAN>

A:Cross-references: EMBL:X72221; NID:g311942; PIDN:CAA51022.1; PID:g311943

C:Superfamily: alpha-s1-casein

Query Match 68.7%; Score 57; DB 2; Length 214;

Best Local Similarity 57.1%; Pred. No. 0.0018;

Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21

||:| ||||| |||

Db 74 QMKAGSSSSSEIVPNSAEQK 94

RESULT 4

casein B precursor - guinea pig

C:Species: Cavia porcellus (guinea pig)

C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 13-Aug-1999

C:Accession: S07130; S27101

R:Hall, L.; Laird, J.E.; Craig, R.K.

Biochem. J. 222, 561-570, 1984

A:Title: Nucleotide sequence determination of guinea-pig casein B mRNA reveals homology

A:Reference number: S07130; MUID:85022410; PMID:6548375

A:Accession: S07130

A:Molecule type: mRNA

A:Residues: 1-198 <HAL1>

A:Cross-references: EMBL:X00938; NID:g49574; PIDN:CAA25452.1; PID:g757817

A:Accession: S27101

A:Molecule type: protein

A:Residues: 16-31 <HAL2>

C:Superfamily: alpha-s1-casein

C:Keywords: milk; phosphoprotein

F:1-15/Domain: signal sequence #status predicted <SIG>

F:16-198/Product: casein B #status experimental <MAT>

Query Match 66.3%; Score 55; DB 2; Length 198;

Best Local Similarity 52.4%; Pred. No. 0.0041;

Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21

||:| ||||| |||

Db 76 QREASSSSSEIVPKNTEQK 96

RESULT 5

alpha-s1-casein - goat

C:Species: Capra aegagrus hircus (domestic goat)

C:Date: 30-Sep-1991 #sequence_revision 26-Jul-1996 #text_change 13-Aug-1999

C:Accession: I46167; S06102; S15489

R:Leroux, C.; Mazure, N.; Martin, P.

J. Biol. Chem. 267, 6147-6157, 1992

A:Title: Mutations away from splice site recognition sequences might cis-modulate altern

A:Reference number: A42158; MUID:92210588; PMID:1372900

A:Accession: I46167

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-214 <LE2>

A:Cross-references: EMBL:X59836; NID:g953; PIDN:CAA42496.1; PID:g954

R:Brignon, G.; Mahe, M.F.; Grosclaude, F.; Ribadeau-Dumas, B.

Protein Seq. Data Anal. 2, 181-188, 1989

A:Title: Sequence of caprine alpha(s1)-casein and characterization of those of its genet

A:Reference number: S06102; MUID:89331508; PMID:2755948

A:Accession: S06102

A:Molecule type: protein

A:Residues: 16-214 <BRI>

A:Note: 23-Ile, 31-Pro, 90-Ser, 92-Glu, 115-Lys and 210-Ala were also found

C:Superfamily: alpha-s1-casein

C:Keywords: phosphoprotein

F:61,63,79,80,81,82,83,90/Binding site: phosphate (Ser) (covalent) #status experimental

F:130/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 65.1%; Score 54; DB 2; Length 214;

Best Local Similarity 52.4%; Pred. No. 0.0071;

Matches 11; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21

||:| ||||| |||

Db 74 QMKAGSSSSSEIVPNSAQK 94

RESULT 6

JC1133

alpha-s1-casein precursor - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 13-Aug-1999

C:Accession: JC1133; S02042

R:Jolivet, G.; Devigny, E.; Fontaine, M.L.; Houdebine, L.M.

Gene 113, 257-262, 1992

A:Title: Structure of the gene encoding rabbit alpha-s1-casein.

A:Reference number: JC1133; MUID:92241677; PMID:1572547

A:Accession: JC1133

A:Molecule type: DNA

A:Residues: 1-215 <JOL>

A:Cross-references: GB:W77195

R:Devigny, E.; Schaefer, E.; Jolivet, G.; Fontaine, M.L.; Kraehenbuhl, J.P.; Houdebine

Nucleic Acids Res. 16, 11813, 1988

A:Title: Sequence of the rabbit alpha(s1)-casein cDNA.

A:Reference number: S02042; MUID:89098329; PMID:3211752

A:Accession: S02042

A:Molecule type: mRNA

A:Residues: 1-215 <DEV>

A:Cross-references: EMBL:X13042; NID:gl513; PIDN:CAA31448.1; PID:gl514

C:Genetics:

A:Gene: Aslca

A:Introns: 17/3; 28/3; 35/3; 49/3; 57/3; 65/3; 79/3; 87/3; 95/3; 105/3; 113/3; 128/3

A:Superfamily: alpha-s1-casein

C:Keywords: milk; phosphoprotein

F:1-15/Domain: signal sequence #status predicted <SIG>

F:16-215/Product: alpha-s1-casein #status predicted <MAT>

Query Match 53.0%; Score 44; DB 2; Length 215;

Best Local Similarity 42.9%; Pred. No. 0.63;

Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21

||:| ||||| |||

Db 95 ETEASSSSSEIVPSSTQK 115

RESULT 7

T23200

hypothetical protein K01G5.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000

C:Accession: T23200

R:Wall, M.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z19706

A:Accession: T23200

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-814 <WIL>

A:Cross-references: EMBL:292803; PIDN:CAB07245.1; GSPDB:GN00021; CESP:K01G5.6

A:Experimental source: clone K01G5

C:Genetics:

A:Gene: CESP:K01G5.6

A:Map position: 3

A:Introns: 7/1; 52/2; 78/2; 112/1; 156/2; 216/3; 702/3; 753/2

C:Superfamily: Caenorhabditis elegans hypothetical protein K01G5.6

Query Match 53.0%; Score 44; DB 2; Length 814;

Best Local Similarity 42.1%; Pred. No. 2.8;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

QY 1 OMEAEIXXXEIVPNXVEQ 19
   | | | | | : | : | | | |
Db 190 OSEKWLNVFORVIFNLVE 208

RESULT 8
T19384
FWRFamide-like peptide 4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T19384; T42411
R:Berk, M.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z19117
A:Accession: T19384
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-93 <WIL>
A:Cross-references: EMBL:Z48543; PIDN:CAA88434.1; GSPDB:GN00020; CBSP:C18D1.3
A:Experimental source: clone C18D1
R:Li, C.; Nelson, L.S.; Memmott, J.M.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z22162
A:Accession: T42411
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-93 <LIC>
A:Cross-references: EMBL:AF042390; PIDN:AAC08941.1
C:Genetics:
A:Gene: flp-4; C18D1.3
A:Map position: 2
A:Introns: 50/3; 69/2

Query Match 49.4%; Score 41; DB 2; Length 93;
Best Local Similarity 33.3%; Pred. No. 0.94;
Matches 7; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 OMEAEIXXXEIVPNXVEQ 21
   : | | | | : | : | | | |
Db 37 EQEKNIASPDLIPEIVREQ 57

RESULT 9
A36060
alpha-casein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 16-Nov-1990 #sequence_revision 16-Nov-1990 #text_change 13-Aug-1999
C:Accession: A36060
R:Grusby, M.J.; Mitchell, S.C.; Nabavi, N.; Glincher, L.H.
Proc. Natl. Acad. Sci. U.S.A. 87, 6897-6901, 1990
A:Title: Casein expression in cytotoxic T lymphocytes.
A:Reference number: A36060; MUID:90370891; PMID:2395885
A:Accession: A36060
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-313 <GRU>
A:Cross-references: GB:M36780; NID:g191574; PIDN:AAA37142.1; PID:g191575
C:Superfamily: alpha-s1-casein
C:keywords: phosphoprotein

Query Match 47.0%; Score 39; DB 2; Length 313;
Best Local Similarity 63.6%; Pred. No. 8.9;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 11 EEIVPNXVEQ 21
   | | | | | | | | | |
Db 100 EEAFINITEQK 110

RESULT 10
C70205
hypothetical protein BB0844 - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)

```

```

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: C70205
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; W.
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
C:Accession: C70205
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-323 <KLE>
A:Cross-references: GB:AE001183; GB:AE000783; NID:g2688786; PIDN:AAC67199.1; PID:g268
A:Experimental source: strain B31

Query Match 47.0%; Score 39; DB 2; Length 323;
Best Local Similarity 40.0%; Pred. No. 9.3;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 MEAEIXXXEIVPNXVEQ 21
   : | | | | | | | | | |
Db 84 LQEEIEIKKEELVPNTDEK 103

RESULT 11
T20355
hypothetical protein D2030.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C:Accession: T20355
R:Wilkinson, J.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19261
A:Accession: T20355
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-508 <WIL>
A:Cross-references: EMBL:Z73906; PIDN:CAA98117.1; GSPDB:GN00019; CBSP:D2030.3
A:Experimental source: clone D2030
C:Genetics:
A:Gene: CESP:D2030.3
A:Map position: 1
A:Introns: 21/3; 160/1; 238/3; 328/3; 391/3
C:Superfamily: Caenorhabditis elegans hypothetical protein D2030.3

Query Match 47.0%; Score 39; DB 2; Length 508;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 EAEXIXXXEIVPNXVEQ 21
   | | | | | | | | | |
Db 432 EMETEDKEEIVPVVVEKK 450

RESULT 12
T10627
hypothetical protein T13K14.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
C:Accession: T10627
R:Bevan, M.; Pohl, T.; Weizensegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lem
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16991
A:Accession: T10627
A:Molecule type: DNA
A:Residues: 1-1396 <BEV>
A:Cross-references: EMBL:AL080282; GSPDB:GN00062; ATSP:T13K14.10
A:Experimental source: cultivar Columbia; BAC clone T13K14
C:Genetics:
A:Gene: ATSP:T13K14.10
A:Map position: 4
A:Introns: 145/1; 171/2; 196/1; 242/3; 269/3; 298/3; 347/3; 374/3; 414/3; 442/3; 505/

```

us-09-380-738a-1.rpr

Wed Feb 12 11:35:21 2003

C;Date: 19-Mar-1993 #sequence_revision 19-Mar-1993 #text_change 24-Sep-1998

C;Accession: A37351

R;Paine-Saunders, S.; Fristrom, D.; Fristrom, J.W.

Dev. Biol. 140, 337-351, 1990

A;Title: The Drosophila IMP-E2 gene encodes an apically secreted protein expressed du

A;Reference number: A37351; MUID:90323384; PMID:2115480

A;Accession: A37351

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-466 <PAI>

A;Cross-references: GB:M55099; NID:g705411; PID:g157752

C;Genetics:

A;Gene: FlyBase:ImpE2

A;Cross-references: FlyBase:FBgn0001254

Query Match 45.8%; Score 38; DB 2; Length 466;

Best Local Similarity 42.1%; Pred. No. 22;

Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 EAEXIXXXEIVPNXVEQK 21

||| : : : |||

Db 171 EETVVVPAEELSPPVEQE 189

Search completed: February 11, 2003, 18:21:12

Job time : 20.4021 secs

131/3; 1139/3; 1163/3; 1194/3; 1226/3; 1259/3; 1290/3; 1344/3

Query Match 47.0%; Score 39; DB 2; Length 1396;

Best Local Similarity 38.1%; Pred. No. 48;

Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21

::||| : : : |||

Db 913 KVEAEALLASEKLIVPLNKK 933

RESULT 13

S66805

hypothetical protein YOL109w - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein O0738

C;Species: Saccharomyces cerevisiae

C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 06-Feb-1998

C;Accession: S66805

R;Durand, P.; Hilger, F.; Portetelle, D.; Vandenbol, M.

submitted to the Protein Sequence Database, July 1996

A;Reference number: S66791

A;Accession: S66805

A;Molecule type: DNA

A;Residues: 1-113 <DUR>

A;Cross-references: EMBL:Z74851; NID:g1419974; PID:e251901; PID:g1419975; MIPS:YOL109w

A;Experimental source: strain S288C

C;Genetics:

A;Gene: SGP:ZEOL

A;Cross-references: SGD:S0005469; MIPS:YOL109w

A;Map position: 15L

Query Match 45.8%; Score 38; DB 2; Length 113;

Best Local Similarity 42.1%; Pred. No. 4.5;

Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 EAEXIXXXEIVPNXVEQK 21

||| : : : |||

Db 51 EAEQVKKEQNADGVEQK 69

RESULT 14

JC6164

circumsporozoite protein - Plasmodium gallinaceum

C;Species: Plasmodium gallinaceum

C;Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 21-Jul-2000

C;Accession: JC6164

R;McCutchan, T.F.; Kissinger, J.C.; Touray, M.G.; Rogers, M.J.; Li, J.; Sullivan, M.; Br

Proc. Natl. Acad. Sci. U.S.A. 93, 11889-11894, 1996

A;Title: Comparison of circumsporozoite proteins from avian and mammalian malarial: Biol

A;Reference number: JC6164; MUID:97030292; PMID:8876233

A;Accession: JC6164

A;Molecule type: DNA

A;Residues: 1-388 <MCC>

A;Cross-references: GB:U65959; NID:g1663571; PID:AC47344.1; PID:g1663572

C;Comment: This protein has an important immunogenicity.

C;Genetics:

A;Gene: CS

C;Superfamily: circumsporozoite protein; thrombospondin type 1 repeat homology

F:312-365/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 45.8%; Score 38; DB 2; Length 388;

Best Local Similarity 72.7%; Pred. No. 18;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 EEIVPNXVEQK 21

||| || |||

Db 347 EETPNDEVK 357

RESULT 15

A37351

IMP-E2 protein precursor - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

us-09-380-738a-1.rsp

Wed Feb 12 11:35:21 2003

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: February 11, 2003, 18:07:35 ; Search time 9.30928 Seconds
(without alignments)
93.563 Million cell updates/sec

Title: US-09-380-738a-1
Perfect score: 83
Sequence: 1 QMEAEIXXXEIVPNXVEQK 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	94.0	214	CAS1_BOVIN	P02662 bos taurus
2	57	68.7	206	CAS1_SHEEP	P04653 ovis aries
3	57	68.7	214	CAS1_CAPI	P18626 capra hircu
4	55	66.3	198	CAS1_CAVPO	P04656 cavia porce
5	44	53.0	215	CAS1_RABIT	P09115 oryctolagus
6	44	53.0	814	EXT2_CAEEL	P01705 caenorhabdi
7	39	47.0	313	CAS1_MOUSE	P19228 mus musculu
8	38	45.8	240	CLP2_SYNP7	P29681 drosophila
9	38	45.8	466	IMP2_DROME	P19228 mus musculu
10	38	45.8	1021	MAPA_MOUSE	P19228 mus musculu
11	37	44.6	739	SYG_HUMAN	P41250 homo sapien
12	36	43.4	206	CAS1_PIG	P39035 sus scrofa
13	36	43.4	840	MUTS_LACLA	Q9cdk9 lactococcus
14	36	43.4	1075	PST2_SCHPO	O13919 schizosacch
15	36	43.4	1338	VGR1_HUMAN	P17948 h vascular
16	35.5	42.8	3396	POLG_DENIS	P33478 d genome po
17	35	42.2	295	VIB6_AGR75	P17796 agrobacteri
18	35	42.2	388	REF_PIRAE	Q8zu81 pyrobaculum
19	35	42.2	575	BCEH_RHOCA	P26168 rhodobacter
20	35	42.2	658	TRK1_STRPN	P22976 streptococc
21	35	42.2	853	YCG1_YEAST	P25588 saccharomyc
22	35	42.2	1462	NKCR_HUMAN	P30414 homo sapien
23	34	41.0	168	ARPL_CAEEL	Q9n3b0 caenorhabdi
24	34	41.0	294	PUR7_THEVO	Q978v1 thermoplasm
25	34	41.0	358	IDPA_MFCGE	P47516 mycoplasma
26	34	41.0	379	IDHB_CAEEL	Q93353 caenorhabdi
27	34	41.0	482	DBP5_YEAST	P20449 saccharomyc
28	34	41.0	499	TRXB_MOUSE	Q9jmm6 mus musculu
29	34	41.0	500	CACT_DROME	Q03017 drosophila
30	33	39.8	353	DPOA_SULTO	Q974t8 sulfolobus
31	33	39.8	428	HEMY_HAEIN	P44772 haemophilus
32	33	39.8	458	YKH3_CAEEL	P34271 caenorhabdi
33	33	39.8	507	ATPO_BRANA	P22201 brassica na

34	33	39.8	507	1	ATPO_RAPSA	P23413 raphanus sa
35	33	39.8	559	1	DNL1_PYRKO	O9hhc4 pyrococcus
36	33	39.8	666	1	TKT_BACHD	O9kad7 bacillus ha
37	33	39.8	784	1	DPO2_AERPE	O93746 aeropyrum p
38	33	39.8	832	1	YFC4_YEAST	P43572 saccharomyc
39	33	39.8	832	1	VGR1_MOUSE	P35969 mus musculu
40	33	39.8	1336	1	VGR1_RAT	P53767 rattus norv
41	33	39.8	1342	1	Z335_HUMAN	Q9h4z2 homo sapien
42	33	39.8	1569	1	GLI3_XENLA	Q91660 xenopus lae
43	33	39.8	1781	1	AKAC_HUMAN	Q02952 homo sapien
44	33	39.8	2774	1	MAPA_RAT	P34926 rattus norv
45	32	38.6	110	1	YCKD_BACSU	P42402 bacillus su

ALIGNMENTS

RESULT 1					
ID	CAS1_BOVIN	STANDARD:	PRT:	214 AA.	
AC	P02662; Q28048;				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	01-NOV-1990 (Rel. 16, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Alpha-S1 casein precursor.				
GN	CSN1S1.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Bovinae; Bos.				
OX	NCBI_TaxID=9913;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=84221403; PubMed=6328443;				
RA	Stewart A.F., Willis I.M., Mackinlay A.G.;				
RT	"Nucleotide sequences of bovine alpha S1- and kappa-casein cDNAs.;"				
RL	Nucleic Acids Res. 12:3895-3907(1984).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Nagao M., Maki M., Sasaki R., Chiba R.;				
RT	"Isolation and sequence analysis of bovine alpha-S1-casein cDNA clone.;"				
RL	Agric. Biol. Chem. 48:1663-1667(1984).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=87049835; PubMed=3022833;				
RA	Gorodetskii S.I., Zakhar'ev V.M., Kyarshulite D.R., Kapelinskaya T.V.;				
RT	Skrvabin K.G.;				
RL	"cDNA of cattle alpha S1-casein: cloning and nucleotide sequence.;"				
RN	Biokhimiia 51:1641-1648(1986).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92051301; PubMed=1658736;				
RA	Koczan D., Hobom G., Seyfert H.M.;				
RT	"Genomic organization of the bovine alpha-S1 casein gene.;"				
RL	Nucleic Acids Res. 19:5591-5596(1991).				
RN	[5]				
RP	SEQUENCE OF 55-130 FROM N.A.				
RX	MEDLINE=83182023; PubMed=6897774;				
RA	Willis I.M., Stewart A.F., Caputo A., Thompson A.R., McKinlay A.G.;				
RT	"Construction and identification by partial nucleotide sequence analysis of bovine casein and beta-lactoglobulin cDNA clones.;"				
RL	DNA 1:375-386(1982).				
RN	[6]				
RP	SEQUENCE OF 122-214 FROM N.A.				
RX	MEDLINE=85178933; PubMed=3838718;				
RA	Kiarshulite D.R., Zakhar'ev V.M., Gorodetskii S.I.;				
RT	"Nucleotide sequence of the 3'-nontranslated region of the mRNA of alpha S1-casein in cows.;"				
RL	Dokl. Akad. Nauk SSSR 280:1433-1437(1985).				
RN	[7]				
RP	SEQUENCE OF 164-214 FROM N.A.				
RX	MEDLINE=94154154; PubMed=1343827;				


```

DR EMBL; X03237; CAA26982.1; -.
DR PIR; A25069; KASHS1.
DR InterPro; IPR001588; Casein.
DR pfam; PF00363; caseins; 1.
DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
DR MILK; Phosphorylation; Signal.
FT SIGNAL 1 15
FT CHAIN 16 206
FT DOMAIN 105 111
FT MOD_RES 79 79
FT MOD_RES 80 80
FT MOD_RES 81 81
FT MOD_RES 82 82
FT MOD_RES 83 83
FT MOD_RES 84 84
FT SEQUENCE 206 AA; 23303 MW; 26955ED115A5852B CRC64;
Query Match 68.7%; Score 57; DB 1; Length 206;
Best Local Similarity 57.1%; Pred. No. 0.0029;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
Oy 1 QMEAEIXXXEIEIVPNXVEQK 21
||:| |||||| |||
Db 74 QMKAGSSSSSEIEIVPNSAEQK 94

RESULT 3
CASL_CAPHI STANDARD; PRT; 214 AA.
AC P18626;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Alpha-S1 casein precursor (Alpha-S1-CN) (Variants A, B, C, D, E
DE and F).
DE CSNLS1.
GN Capra hircus (Goat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A. (VARIANT A).
RX MEDLINE=92210388; PubMed=1372900;
RA Leroux C., Mazure N., Martin P.;
RT "Mutations away from splice site recognition sequences might cis-
RT modulate alternative splicing of goat alpha s1-casein transcripts.
RT Structural organization of the relevant gene.";
RL J. Biol. Chem. 267:6147-6157(1992).
RN [2]
RP SEQUENCE FROM N.A. (VARIANT E).
RC TISSUE=Mammary gland;
RX MEDLINE=95011611; PubMed=7926797;
RA Perez M.J., Leroux C., Bonastre A.S., Martin P.;
RT "occurrence of a LINE sequence in the 3' UTR of the goat alpha
RT s1-casein E-encoding allele associated with reduced protein synthesis
RT level.";
RL Gene 147:179-187(1994).
RN [3]
RP SEQUENCE OF 16-214 (VARIANTS D; E AND F).
RX MEDLINE=91031451; PubMed=2226443;
RA Brignon G., Mahe M.-F., Ribadeau-Dumas B., Mercier J.-C.,
RA Grosclaude F.;
RT "Two of the three genetic variants of goat alpha s1-casein which are
RT synthesized at a reduced level have an internal deletion possibly due
RT to altered RNA splicing.";
RL Eur. J. Biochem. 193:237-241(1990).
RN [4]
RP SEQUENCE OF 16-214 (VARIANTS A; B AND C).
RX MEDLINE=89331508; PubMed=2755948;
RA Brignon G., Mahe M.-F., Grosclaude F., Ribadeau-Dumas B.;
RT "Sequence of caprine alpha s1-casein and characterization of those of
RT its genetic variants which are synthesized at a high level. alpha s1-

```

```

RT conservation of the non-coding regions of the mRNA. ";
RL Biochem. J. 222:561-570(1984).
CC -!- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT
CC CALCIUM PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: BELONGS TO THE ALPHA-CASEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X00938; CAA25452.1; -
DR PIR; S07130; S07130.
DR InterPro; IPR001588; Casein.
DR Pfam; PF00363; caseins; 1.
DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
KW Milk; Phosphorylation; Signal.
FT SIGNAL 1 15
FT CHAIN 16 198
FT MOD_RES 34 34 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 35 35 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 36 36 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 37 37 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 38 38 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 39 39 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 40 40 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 80 80 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 81 81 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 83 83 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 84 84 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 85 85 PHOSPHORYLATION (POTENTIAL).
SQ SEQUENCE 198 AA; 23140 MW; C63A72286A003EBF CRC64;

Query Match 66.3%; Score 55; DB 1; Length 198;
Best Local Similarity 52.4%; Pred. No. 0.0064;
Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
I ||| ||||| |||
Db 76 QREASSISSEIVPKNTEQK 96

RESULT 5
CASI_RABIT STANDARD; PRT; 215 AA.
AC P09115;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Alpha casein precursor (Alpha-s1 casein).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Mammary gland;
RX MEDLINE=89098329; PubMed=3211752;
RA Devinoy E., Schaefer E., Jolivet G., Fontaine M.L., Kraehenbuhl J.P.,
RA Houdebine L.M.;
RT "Sequence of the rabbit alpha S1-casein cDNA.";
RL Nucleic Acids Res. 16:11813-11813(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92241677; PubMed=1572547;
RA Jolivet G., Devinoy E., Fontaine M.L., Houdebine L.M.;
RT "Structure of the gene encoding rabbit alpha s1-casein.";
RL Gene 113:257-262(1992).

```

```

CC -!- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT
CC CALCIUM PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: BELONGS TO THE ALPHA-CASEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13042; CAA31448.1; -
DR PIR; S02042; S02042.
DR PIR; JC1133; JC1133.
DR InterPro; IPR001588; Casein.
DR Pfam; PF00363; caseins; 1.
DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
KW Milk; Phosphorylation; Signal.
FT SIGNAL 1 15
FT CHAIN 16 215
FT MOD_RES 101 101 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 102 102 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 103 103 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 104 104 PHOSPHORYLATION (POTENTIAL).
SQ SEQUENCE 215 AA; 25528 MW; 40737A027636B255 CRC64;

Query Match 53.0%; Score 44; DB 1; Length 215;
Best Local Similarity 42.9%; Pred. No. 0.67;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
I ||| ||||| |||
Db 95 ETEASISSSEIVPSSTKQK 115

RESULT 6
EXT2_CAEL STANDARD; PRT; 814 AA.
AC Q01705; Q17920;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Exostosin-2 (EC 2.4.1.-) (Multiple exostos homolog 2).
GN RIB-2 OR K01G5.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97264342; PubMed=9110175;
RA Clines G.A., Ashley J.A., Shah S., Lovett M.;
RT "The structure of the human multiple exostos 2 gene and
RT characterization of homologs in mouse and Caenorhabditis elegans.";
RL Genome Res. 7:359-367(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=121265000; PubMed=11121397;
RA Kitagawa H., Egusa N., Tamura J.I., Kusche-Gullberg M., Lindahl U.,
RA Sugihara K.;
RT "rib-2, a Caenorhabditis elegans homolog of the human tumor suppressor
RT EXT genes encodes a novel alpha1,4-N-acetylglucosaminyltransferase
RT involved in the biosynthetic initiation and elongation of heparan
RT sulfate.";
RL J. Biol. Chem. 276:4834-4838(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wall M.;

```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 -1- FUNCTION: Probable alpha,4-N-acetylglucosaminyltransferase required for the biosynthesis of heparan-sulfate (By similarity).
 -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic reticulum (By similarity).
 -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 47.
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U94835; AAC47510.1; -
 EMBL: AB077851; BAB83878.1; -
 EMBL: 292803; CAB07245.1; -
 WormPep: K01G5.6; C016196.
 InterPro: IPR004263; Exostosin.
 Pfam: PF03016; Exostosin 1.
 Transferrase; Glycosyltransferase; Endoplasmic reticulum;
 Transmembrane; Signal-anchor; Glycoprotein.
 DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 15 35 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).
 DOMAIN 36 814 LUMENAL (POTENTIAL).
 FT CARBOHYD 36 36 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 492 492 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 685 685 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 604 604 T -> S (IN REF. 1).
 FT SEQUENCE 814 AA; 94196 MW; 38FA58C5EB17DB16 CRC64;
 Query Match 53.0%; Score 44; DB 1; Length 814;
 Best Local Similarity 42.1%; Pred. No. 2.3;
 Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 QMEAEIXXXEIVPNXVE 19
 Db 190 QSEKEWLVQEVIPNLVE 208
 ||| : :||| |||
 190 QSEKEWLVQEVIPNLVE 208

RESULT 7
 CASI_MOUSE STANDARD; PRT; 313 AA.
 ID CASI_MOUSE
 AC F19228;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DE Alpha casein precursor.
 DE CSN1 OR CSNA.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90370891; PubMed=2395885;
 RA Grusby M.J., Mitchell S.C., Nabavi N., Glimcher L.H.;
 RT "Casein expression in cytotoxic T lymphocytes."
 RL Proc. Natl. Acad. Sci. U.S.A. 87:6897-6901(1990).
 CC -1- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT CALCIUM PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
 CC -1- SIMILARITY: BELONGS TO THE ALPHA-CASEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

EMBL: M36780; AAA37142.1; -
 EIR; A36060; A36060.
 MGD; MGI:88540; CsnA.
 InterPro: IPR001588; Casein.
 Pfam: PF00363; caseins; 1.
 PROSITE: PS00306; CASEIN ALPHA_BETA; 1.
 Milk; Phosphorylation; Repeat; Signal.
 SIGNAL 1 15
 FT CHAIN 16 313 ALPHA CASEIN.
 FT MOD_RES 34 34 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 35 35 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 36 36 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 93 93 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 94 94 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 95 95 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 96 96 PHOSPHORYLATION (POTENTIAL).
 FT DOMAIN 135 224 15 X 6 AA TANDEM REPEATS.
 FT REPEAT 135 140 1.
 FT REPEAT 141 146 2.
 FT REPEAT 147 152 3.
 FT REPEAT 153 158 4.
 FT REPEAT 159 164 5.
 FT REPEAT 165 170 6.
 FT REPEAT 171 176 7.
 FT REPEAT 177 182 8.
 FT REPEAT 183 188 9.
 FT REPEAT 189 194 10.
 FT REPEAT 195 200 11.
 FT REPEAT 201 206 12.
 FT REPEAT 207 212 13.
 FT REPEAT 213 218 14.
 FT REPEAT 219 224 15.
 SQ SEQUENCE 313 AA; 35602 MW; 7006C03C59409D70 CRC64;

Query Match 47.0%; Score 39; DB 1; Length 313;
 Best Local Similarity 63.6%; Pred. No. 7.6;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 11 EETVFNXVEOK 21
 Db 100 EEAIPTITEOK 110
 ||| : :||| |||
 100 EEAIPTITEOK 110

RESULT 8
 CLP2_SYN7 STANDARD; PRT; 240 AA.
 ID CLP2_SYN7
 AC O34125;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE ATP-dependent Clp protease proteolytic subunit 2 (EC 3.4.21.92)
 DE (Endopeptidase Clp 2).
 DE CLP2.
 GN Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
 OS Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
 OX NCBI_TaxID=1140;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Clarke A.K., Schelin J.;
 RT "The clp2/X operon from the cyanobacterium Synecococcus sp. strain PCC 7942."
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Cleaves peptides in various proteins in a process that requires ATP hydrolysis. Has a chymotrypsin-like activity. Plays a major role in the degradation of misfolded proteins (By similarity).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins to small peptides in

us-09-380-738a-1.rsp

Wed, Feb 12 11:35:21 2003

```

Query Match          45.8%; Score 38; DB 1; Length 466;
Best Local Similarity 42.1%; Pred. No. 17;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 EAEXIXXXEIVPNXVEQK 21
    | : ||| : ||| :
Db 171 EETVVVPAEELSPVPEQE 189

RESULT 10
MAPA_MOUSE STANDARD; PRT: 1021 AA.
ID MAPA_MOUSE
AC Q9QYR6; Q9QZH9; Q9QZIO; Q9QZIL;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein 1A (MAP 1A) (Fragment).
DE MAP1A OR MTAP1.
GN Mus musculus (Mouse).
OS Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_TaxID=10090;
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RC TISSUE=Brain;
RC MEDLINE=21210968; PubMed=11311937;
RA Nakayama A., Odajima T., Murakami H., Mori N., Takahashi M.;
RT "Characterization of two promoters that regulate alternative
RT transcripts in the microtubule-associated protein (MAP) 1A gene.";
RL Biochim. Biophys. Acta 1518:260-266(2001).
CC -1- FUNCTION: Structural protein involved in the filamentous cross-
CC bridging between microtubules and other skeletal elements.
CC -1- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate
CC with MAP1A and MAP1B proteins.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC short form; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Both isoforms highly expressed in brain, and
CC to a lesser extent in embryo. The short isoform is also expressed
CC at a low level in other tissues including heart and muscle.
CC -1- DOMAIN: The basic region containing the repeats may be responsible
CC for the binding of MAP1A to microtubules.
CC -1- PTM: Various serine residues may be phosphorylated by cAMP kinase
CC (BV similarity).
CC -1- SIMILARITY: TO MAP1B.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF182211; AAF06164.1; JOINED.
CC EMBL; AF182208; AAF06164.1; JOINED.
CC EMBL; AF182209; AAF06164.1; JOINED.
CC EMBL; AF182211; AAF06163.1;
CC EMBL; AF182213; AAD55790.1;
CC EMBL; AF182212; AAD55789.1;
CC MGD; MGI:97174; Mtap1.
CC Microtubules; Repeat; Phosphorylation; Alternative splicing.
KW LYS-RICH (BASIC).
FT FT 533 720
FT FT 560 765
FT FT 560 562
FT FT 560 641
FT FT 639 641
FT FT 644 646
FT FT 648 650
FT FT 651 653
FT FT 655 657
FT FT 660 662
FT FT 660 666
FT FT 664 666

us-09-380-738a-1.rsp
Query Match          45.8%; Score 38; DB 1; Length 1021;
Best Local Similarity 35.0%; Pred. No. 34;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 MEAEIXXXEIVPNXVEQK 21
    | : ||| : ||| :
Db 807 LEQEQVEREKEVVPDPEDK 826

RESULT 11
SYG_HUMAN STANDARD; PRT: 739 AA.
ID SYG_HUMAN
AC P41250; Q969Y1;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glycyl-tRNA synthetase (EC 6.1.1.14) (Glycine--tRNA ligase) (GlyRS).
DE GARS.
GN Homo sapiens (Human).
OS Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
OX [1]
RN [1]
RP SEQUENCE FROM N.A., PubMed=7962006;
RX MEDLINE=95050870; PubMed=7962006;
RA Shiba K., Schimmel P., Motegi H., Noda T.;
RT "Human glycyl-tRNA synthetase. Wide divergence of primary structure
RT from bacterial counterpart and species-specific aminoacylation.";
RT J. Biol. Chem. 269:30049-30055(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95273165; PubMed=7753621;
RA Williams J.H., Osvath S.R., Khong T.-F., Pearce M.J., Power D.A.;
RT "Cloning, sequencing and bacterial expression of human glycine tRNA
RT synthetase.";
RT Nucleic Acids Res. 23:1307-1310(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Eye, and Muscle;
RA Strausberg P.;
RN Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 3-739 FROM N.A.
RX MEDLINE=95050687; PubMed=7961834;
RA Ge Q., Trieu E.P., Taroff I.N.;
RT "Primary structure and functional expression of human glycyl-tRNA
RT synthetase, an autoantigen in myositis.";
RT J. Biol. Chem. 269:28790-28797(1994).
RN [5]
RP SEQUENCE OF 348-739 FROM N.A.
RA Andrews S., Langston Y., Stoneking T., Maupin R.;
RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + glycine + tRNA(Gly) = AMP + diphosphate
CC + glycyl-tRNA(Gly).
CC -1- SUBUNIT: Homodimer.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 WHEP-TRS DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC

```


DR EMBL: D30658; BAA06338.1; .
 DR EMBL: U09510; AAA86443.1; ALT_INIT.
 DR EMBL: BC007722; AAH07722.1; .
 DR EMBL: BC007755; AAH07755.1; .
 DR EMBL: U09587; AA57001.1; ALT_INIT.
 DR EMBL: AC004976; AAC71652.1; .
 DR HSP: P56206; LATI.
 DR Genew: HGNC:4162; GARS.
 DR MIM: 600287; .
 DR InterPro: IPR002106; AALRNA_ligaseII.
 DR InterPro: IPR004154; HGTP_anticon.
 DR InterPro: IPR00738; WHEP-TRS.
 DR InterPro: IPR002314; tRNA-synt_2b.
 DR Pfam: PF00458; WHEP-TRS; 1.
 DR Pfam: PF00587; tRNA-synt_gly.
 DR Pfam: PF03129; tRNA-synt_2b; 1.
 DR Pfam: PF03129; HGTP_anticon; 1.
 DR PRINTS: PR01043; TRNASYNTHGLY.
 DR TIGRFAMs: TIGR00389; glys_dimeric; 1.
 DR PROSITE: PS00862; AA_TRNA_LIGASE-II; 1.
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT DOMAIN 74 119 M -> I (IN REF. 2).
 FT CONFLICT 530 530 WHEP-TRS.
 SQ SEQUENCE 739 AA; 83139 MW; 55DD57119F438E5 CRC64;
 Query Match 44.6%; Score 37; DB 1; Length 739;
 Best Local Similarity 66.7%; Pred. No. 39;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 11 EEIVPNKVE 19
 Db 568 EEVVPNVIE 576
 ||:||||
 RESULT 12
 CASL_PIG
 ID CASL_PIG STANDARD; PRT; 206 AA.
 AC P39035;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DE 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Alpha-S1 casein precursor.
 GN CSN1S1.
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=92367946; PubMed=1503265;
 RA Alexander L.J., Beattie C.W.;
 RT "The sequence of porcine alpha s1-casein cDNA: evidence for protein
 RT variants generated by altered alpha s1-casein splicing.";
 RL Anim. Genet. 23:283-288(1992).
 CC -!- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT
 CC CALCIUM PHOSPHATE.
 CC -!- SURCELLULAR LOCATION: Extracellular.
 CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
 CC -!- SIMILARITY: BELONGS TO THE ALPHA-CASEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions. There are no restrictions on its
 CC modified and this statement is as long as its content is in no way
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: X54973; CAA38717.1; .
 DR PIR: A45661; A45661.
 DR InterPro: IPR001588; Casein.

DR Pfam: PF00363; caseins; 1.
 DR PROSITE: PS00306; CASEIN_ALPHA_BETA; FALSE_NEG.
 KW Milk; Phosphorylation; Signal.
 FT SIGNAL 1 15
 FT CHAIN 16 206
 FT ALPHA-S1 CASEIN.
 SQ SEQUENCE 206 AA; 24148 MW; 58670C0B309A4456 CRC64;

Query Match 43.4%; Score 36; DB 1; Length 206;
 Best Local Similarity 63.6%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 11 EEIVPNKVEOK 21
 Db 83 EEVVGNSAEOK 93
 ||:||||

RESULT 13
 MUTS_LACLA
 ID MUTS_LACLA STANDARD; PRT; 840 AA.
 AC Q9CDK9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA mismatch repair protein muts.
 GN MUTS OR HEXA OR LL2210.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235186; PubMed=1137471;
 RA Bolotin A., Wincker P., Manger S., Jaillon O., Malarne K.,
 RA Weissenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis IL1403";
 RL Genome Res. 11:731-753(2001).
 CC -!- FUNCTION: This protein is involved in the repair of mismatches in
 CC DNA. It is possible that it carries out the mismatch recognition
 CC step. This protein has a weak ATPase activity (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions. There are no restrictions on its
 CC modified and this statement is as long as its content is in no way
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: AE006450; AAK06308.1; .
 DR InterPro: IPR000432; Muts_C.
 DR InterPro: IPR002863; Muts_N.
 DR Pfam: PF00488; Muts_C; 1.
 DR Pfam: PF01624; Muts_N; 2.
 DR ProDom: PD001263; Muts_C; 1.
 DR SMART: SM00534; Mutsag; 1.
 DR SMART: SM00533; Mutsd; 1.
 DR TIGRFAMs: TIGR01070; muts1; 1.
 DR PROSITE: PS00486; DNA_MISMATCH_REPAIR_2; 1.
 KW DNA repair; ATP-binding; DNA-binding; Complete proteome.
 FT NP_BIND 601 608 ATP (POTENTIAL).
 SQ SEQUENCE 840 AA; 94272 MW; 867E4F82616DC2 CRC64;

Query Match 43.4%; Score 36; DB 1; Length 840;
 Best Local Similarity 40.0%; Pred. No. 66;
 Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 5 EXIXXEEIVPNKVE 19
 Db 575 EAVMGAEVVPNDIE 589
 ||:||||

us-09-380-738a-1.rsp

Wed, Feb 12 11:35:21 2003

```

1 QMEAXLXXXXEIVPNKVEOK 21
: : : : : : : : : : : :
768 RLEQVSIIOQRIIPVPSOK 788

RESULT 15
VGR1_HUMAN STANDARD; PRT; 1338 AA.
ID P1748; P16057; O60722; Q12954;
AC 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Vascular endothelial growth factor receptor 1 precursor (EC 2.7.1.112)
DE (VEGFR-1) (Vascular permeability factor receptor) (Tyrosine-protein
DE kinase receptor FLT) (Flt-1) (Tyrosine-protein kinase FRT) (fms-like
DE tyrosine kinase 1).
DE FLT1 OR FLT OR FRT.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORM FLT1).
TISSUE=Placenta;
MEDLINE=90221591; PubMed=2158038;
Shibuya M., Yamaguchi S., Yamane A., Ikeda T., Tojo A.,
Matsushima H., Sato M.;
"Nucleotide sequence and expression of a novel human receptor-type
tyrosine kinase gene (flt) closely related to the fms family.";
Oncogene 5:519-524(1990).
[2]
SEQUENCE FROM N.A. (ISOFORM FLT1).
TISSUE=Umbilical vein;
Yu Y., Whitney R.G., Sato J.D.;
"Coding region for human VEGF receptor FLT1 (VEGFR-1).";
Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A. (ISOFORM SFLT1), AND SEQUENCE OF N-TERMINUS.
TISSUE=Umbilical vein;
MEDLINE=94068470; PubMed=8248162;
Kendall R.L., Thomas K.A.;
"Identification of vascular endothelial cell growth factor activity by an
endogenously encoded soluble receptor.";
Proc. Natl. Acad. Sci. U.S.A. 90:10705-10709(1993).
[4]
SEQUENCE OF 1018-1058 FROM N.A. (ISOFORM FLT1).
MEDLINE=87307638; PubMed=3040650;
Matsushima H., Yoshida M.C., Sasaki M., Shibuya M.;
"A possible new member of tyrosine kinase family, human fms-like
kinase receptor FLT1, is highly conserved in vertebrates and located on human chromosome
13.";
Jpn. J. Cancer Res. 78:655-661(1987).
[5]
PARTIAL SEQUENCE, PHOSPHORYLATION SITES, AND MUTAGENESIS OF TYR-914;
TYR-1213; TYR-1242; TYR-1327 AND TYR-1333.
PubMed=9722576;
Ito N., Wernstedt C., Engstrom U., Claesson-Welsh L.;
"Identification of vascular endothelial growth factor receptor-1
tyrosine phosphorylation sites and binding of SH2 domain-containing
molecules.";
J. Biol. Chem. 273:23410-23418(1998).
[6]
STRUCTURE BY NMR OF 129-229.
MEDLINE=20013066; PubMed=10543948;
Starovasnik M.A., Christinger H.W., Wiesmann C., Champe M.A.;
"Solution structure of the VEGF-binding domain of Flt-1: comparison
of its free and bound states.";
J. Mol. Biol. 293:531-544(1999).
CC -!- FUNCTION: Receptor for VEGF, VEGFB and bFGF. Has a tyrosine-protein
kinase activity. The VEGF-kinase ligand/receptor signaling system
plays a key role in vascular development and regulation of

```

Query Match 43.4%; Score 36; DB 1; Length 1075;
 Best Local Similarity 33.3%; Pred. No. 83;
 Matches 7; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 OS Alpha sl-casein A short form.
 OC Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95325458; PubMed=7601973;
 RA Ferranti P., Malorni A., Nitti G., Laezza P., Pizzano R., Chianese L.,
 RA Addeo F.;
 RT "Primary structure of ovine alpha sl-caseins: localization of
 RT phosphorylation sites and characterization of genetic variants A, C
 RT and D.";
 RL J. Dairy Res. 62:281-296(1995).
 DR InterPro; IPR001588; Casein.
 DR Pfam; PF00363; caseins; 1.
 SQ SEQUENCE 191 AA; 21739 MW; 37C50FE3A4829ABE CRC64;
 Query Match 68.7%; Score 57; DB 6; Length 191;
 Best Local Similarity 57.1%; Pred. No. 0.0054;
 Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
 QY 1 QMEAEIXXXEIVPNXVEQK 21
 Db 59 QMKAGSSSSSEIVPNSAEQK 79
 RESULT 3
 Q9TS48
 ID Q9TS48 PRELIMINARY; PRT; 199 AA.
 AC Q9TS48
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE Alpha sl-casein A long form.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95325458; PubMed=7601973;
 RA Ferranti P., Malorni A., Nitti G., Laezza P., Pizzano R., Chianese L.,
 RA Addeo F.;
 RT "Primary structure of ovine alpha sl-caseins: localization of
 RT phosphorylation sites and characterization of genetic variants A, C
 RT and D.";
 RL J. Dairy Res. 62:281-296(1995).
 DR InterPro; IPR001588; Casein.
 DR Pfam; PF00363; caseins; 1.
 SQ SEQUENCE 199 AA; 22751 MW; C8D75CB2283E9A5B CRC64;
 Query Match 68.7%; Score 57; DB 6; Length 199;
 Best Local Similarity 57.1%; Pred. No. 0.0056;
 Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
 QY 1 QMEAEIXXXEIVPNXVEQK 21
 Db 59 QMKAGSSSSSEIVPNSAEQK 79
 RESULT 4
 Q8SPR1
 ID Q8SPR1 PRELIMINARY; PRT; 212 AA.
 AC Q8SPR1
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DE Alpha sl casein.

OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lenasi T., Rogelj I., Dovc P.;
 RT "Equus caballus alpha-si-casein (asl-CN) mRNA.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF040862; AAK83668.1;
 SQ SEQUENCE 212 AA; 25305 MW; 578F72EA76E2656E CRC64;
 Query Match 54.2%; Score 45; DB 6; Length 212;
 Best Local Similarity 42.9%; Pred. No. 1.2;
 Matches 9; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
 QY 1 QMEAEIXXXEIVPNXVEQK 21
 Db 79 QOESSSTSSSERVVPINTEQK 99
 RESULT 5
 Q09480
 ID Q09480 PRELIMINARY; PRT; 93 AA.
 AC Q09480;
 DT 01-JAN-1999 (Tremblrel. 09, Created)
 DT 01-JAN-1999 (Tremblrel. 09, Last sequence update)
 DE Hypothetical 10.3 kDa protein Cl8D1.3 in chromosome II.
 GN Cl8D1.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Berks M.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL;
 RA Li C., Nelson L.S., Memmott J.M.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z48543; CAAB88434.1;
 DR EMBL; AF042390; AAC08941.1;
 DR WormPep; Cl8D1.3; CE01503.
 KW Hypothetical protein.
 SQ SEQUENCE 93 AA; 10347 MW; DADC271BC6F34FF2 CRC64;
 Query Match 49.4%; Score 41; DB 5; Length 93;
 Best Local Similarity 33.3%; Pred. No. 3;
 Matches 7; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
 QY 1 QMEAEIXXXEIVPNXVEQK 21
 Db 37 EQEENKIASPDELIPEIVEQK 57
 RESULT 6
 Q99M19
 ID Q99M19 PRELIMINARY; PRT; 292 AA.
 AC Q99M19;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DE Similar to casein alpha.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RA Strausberg R.; to the EMBL/GenBank/DBJ databases.

RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC002101; AA02101.1; -.

DR InterPro; IPR001588; Casein.

DR Pfam; PF00363; caseins; 1.

DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.

SQ SEQUENCE 292 AA; 32988 MW; AF3C9B7A4899BF5 CRC64;

Query Match 47.0%; Score 39; DB 11; Length 292;

Best Local Similarity 63.6%; Pred. No. 23;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 11 EEIVPNXVEQK 21

Db 98 EEAIPTITEQK 108

RESULT 7

Q91VVI PRELIMINARY; PRT; 294 AA.

ID Q91VVI

AC Q91VVI

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Similar to casein alpha.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS; TISSUE=TESTIS

RA Strausberg R.; to the EMBL/GenBank/DBJ databases.

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC008278; AA08278.1; -.

DR InterPro; IPR001588; Casein.

DR Pfam; PF00363; caseins; 1.

DR PROSITE; PS00306; CASEIN_ALPHA_BETA; UNKNOWN_1.

SQ SEQUENCE 294 AA; 33244 MW; 02347DF26D5098A CRC64;

Query Match 47.0%; Score 39; DB 11; Length 294;

Best Local Similarity 63.6%; Pred. No. 23;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 11 EEIVPNXVEQK 21

Db 100 EEAIPTITEQK 110

RESULT 8

Q99JM6 PRELIMINARY; PRT; 305 AA.

ID Q99JM6

AC Q99JM6

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)

DE Similar to casein alpha.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS; TISSUE=TESTIS

RA Strausberg R.; to the EMBL/GenBank/DBJ databases.

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC006024; AA06024.1; -.

DR InterPro; IPR001588; Casein.

DR Pfam; PF00363; caseins; 1.

DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.

SQ SEQUENCE 305 AA; 34797 MW; 04800330BA6C38BF CRC64;

Query Match 47.0%; Score 39; DB 11; Length 305;

Best Local Similarity 63.6%; Pred. No. 23;

QY 11 EEIVPNXVEQK 21

Db 92 EEAIPTITEQK 102

RESULT 9

O51784 PRELIMINARY; PRT; 323 AA.

ID O51784

AC O51784

DT 01-JUN-1998 (TREMBlrel. 06, Created)

DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Hypothetical protein BB0844.

GN BB0844.

OS Borrelia burgdorferi (Lyme disease spirochete).

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.

OX NCBI_TaxID=139;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 35210 / B31;

RX MEDLINE=98065943; PubMed=9403685;

RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,

RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,

RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,

RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,

RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,

RA Utterback T., Wathley L., McDonald L., Artiach P., Bowman C.,

RA Carland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,

RA Smith H.O., Venter J.C.;

RT "Genomic sequence of a Lyme disease spirochete, Borrelia

burgdorferi."

RL Nature 390:580-586(1997).

DR EMBL; A5001183; AAC67199.1; -.

DR TIGR; BB0844;

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 323 AA; 37466 MW; 312C1E2E1E849F5D CRC64;

Query Match 47.0%; Score 39; DB 16; Length 323;

Best Local Similarity 40.0%; Pred. No. 25;

Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 MEAEIXXXEEIVPNXVEQK 21

Db 84 LQEEIETKEELVPTDEEK 103

RESULT 10

P90790 PRELIMINARY; PRT; 508 AA.

ID P90790

AC P90790

DT 01-MAY-1997 (TREMBlrel. 03, Created)

DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE D2030.3 protein.

GN D2030.3

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RA Wilkerson J.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;

RA none;

RT "Genome sequence of the nematode C.elegans: A platform for

Science 282:2012-2018(1998).

RL EMBL; 273906; CAA98117.1; -.

```

SQ SEQUENCE 508 AA; 56589 MW; 812AC8891A578B64 CRC64;
Query Match
Best Local Similarity 47.0%; Score 39; DB 5; Length 508;
Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 EAEXIXXXEIVPNXVEQK 21
   ||| : | : | : | : |
Db 432 EMETEDKEEIVPVVEKR 450

RESULT 11
Q9SUC7
ID Q9SUC7 PRELIMINARY; PRT; 1396 AA.
AC Q9SUC7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Hypothetical 154.2 kDa protein.
GN T13K14.10 OR AT4G20850.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Pohl T., Weizsaecker T., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL080282; CAB45880.1; -
DR EMBL; AL161553; CAB79085.1; -
DR HSP; P00782; 2SPT.
DR MEROPS; S08.090; -
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 3.
DR PRINTS; PR00723; SUBTILISIN
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hypothetical protein.
SQ SEQUENCE 1396 AA; 154174 MW; 113209E42E9DE4E3 CRC64;

Query Match
Best Local Similarity 47.0%; Score 39; DB 10; Length 1396;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
   : ||| : | : | : | : |
Db 913 KVEAEALLASEKLPIAVLNK 933

RESULT 12
Q08245
ID Q08245 PRELIMINARY; PRT; 113 AA.
AC Q08245;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Chromosome XV reading frame ORF YOL109W.
GN ZEO1 OR YOL109W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Durand P., Hilger F., Portetelle D., Vandenbol M.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 274851; CAA99128.1; -
DR SGD; S0005469; ZEO1.
SQ SEQUENCE 113 AA; 12589 MW; 1EA3634565D517FA CRC64;

Query Match
Best Local Similarity 45.8%; Score 38; DB 3; Length 113;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 EAEXIXXXEIVPNXVEQK 21
   ||| : | : | : | : |
Db 51 EAEQVKKEQNIADGVEQK 69

RESULT 13
Q95KZ7
ID Q95KZ7 PRELIMINARY; PRT; 208 AA.
AC Q95KZ7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE AlphaS1-casein.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Milenkovic D., Martin P., Guerin G., Leroux C.;
RT "Horse specific pattern of alphaS1-casein RNA splicing and genomic
RT characterization of the relevant locus";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY049939; AAL05435.1; -
DR InterPro; IPR001588; Casein.
DR PROSITE; PS00306; CASEIN_ALPHA_BETA; UNKNOWN_1.
SQ SEQUENCE 208 AA; 24689 MW; 0CC6C409489C589C CRC64;

Query Match
Best Local Similarity 45.8%; Score 38; DB 6; Length 208;
Matches 7; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 QMEAEIXXXEIVPNXVEQK 21
   | : | : | : | : |
Db 87 QQESSSTSSSEVPVPIATEKR 107

RESULT 14
Q94675
ID Q94675 PRELIMINARY; PRT; 388 AA.
AC Q94675;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Circumsporozoite protein.
GN CSP.
OS Plasmodium gallinaceum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5849;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97030292; PubMed=8876233;
RA McCutchan T.F., Kissinger J.C., Touray M.G., Rogers M.J., Li J.,
RA Sullivan M., Braga E.M., Kretzli A.U., Miller L.H.;
RT "Comparison of circumsporozoite proteins from avian and mammalian
RT malarial: biological and phylogenetic implications";
RL Proc. Natl. Acad. Sci. U.S.A. 93:11889-11894 (1996).
DR EMBL; U05959; AAC47344.1; -

```

us-09-380-738a-1.l.rspt

Wed Feb 12 11:35:21 2003

DR InterPro: IPR003067; Crmsprzoite.

DR InterPro: IPR000884; TSPL.

DR Pfam: PF00090; tsp_1; 1.

DR PRINTS: PR01303; CRMSPRZOITE.

DR SMART: SM00209; TSPL; 1.

SQ SEQUENCE 388 AA; 39812 MW; CF493FD9C6CE6A62 CRC64;

Query Match 45.8%; Score 38; DB 5; Length 388;

Best Local Similarity 72.7%; Pred. No. 46;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 EEIVPNXVEQK 21

DB 347 EEITPNDEVK 357

RESULT 15

O8RWX2 PRELIMINARY; PRT; 504 AA.

ID O8RWX2

AC O8RWX2; 21, Created)

DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)

DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)

DE Hypothetical 57.7 kDa protein.

GN AT2G32170.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Yamada K., Ban H., Chan M.M., Chang C.H., Chang E., Dale J.M.,

RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,

RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,

RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,

RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,

RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,

RA Sakurai T., Satou J.R., Shinn P., Southwick A., Shinozaki K.,

RA Davis R.W., Ecker J.R., Theologis A.;

RT "Arabidopsis Full Length cDNA Clones."

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY091048; AAM13869.1; -.

KW Hypothetical protein.

SQ SEQUENCE 504 AA; 57711 MW; DFE306FD368EC59C CRC64;

Query Match 45.8%; Score 38; DB 10; Length 504;

Best Local Similarity 42.1%; Pred. No. 60;

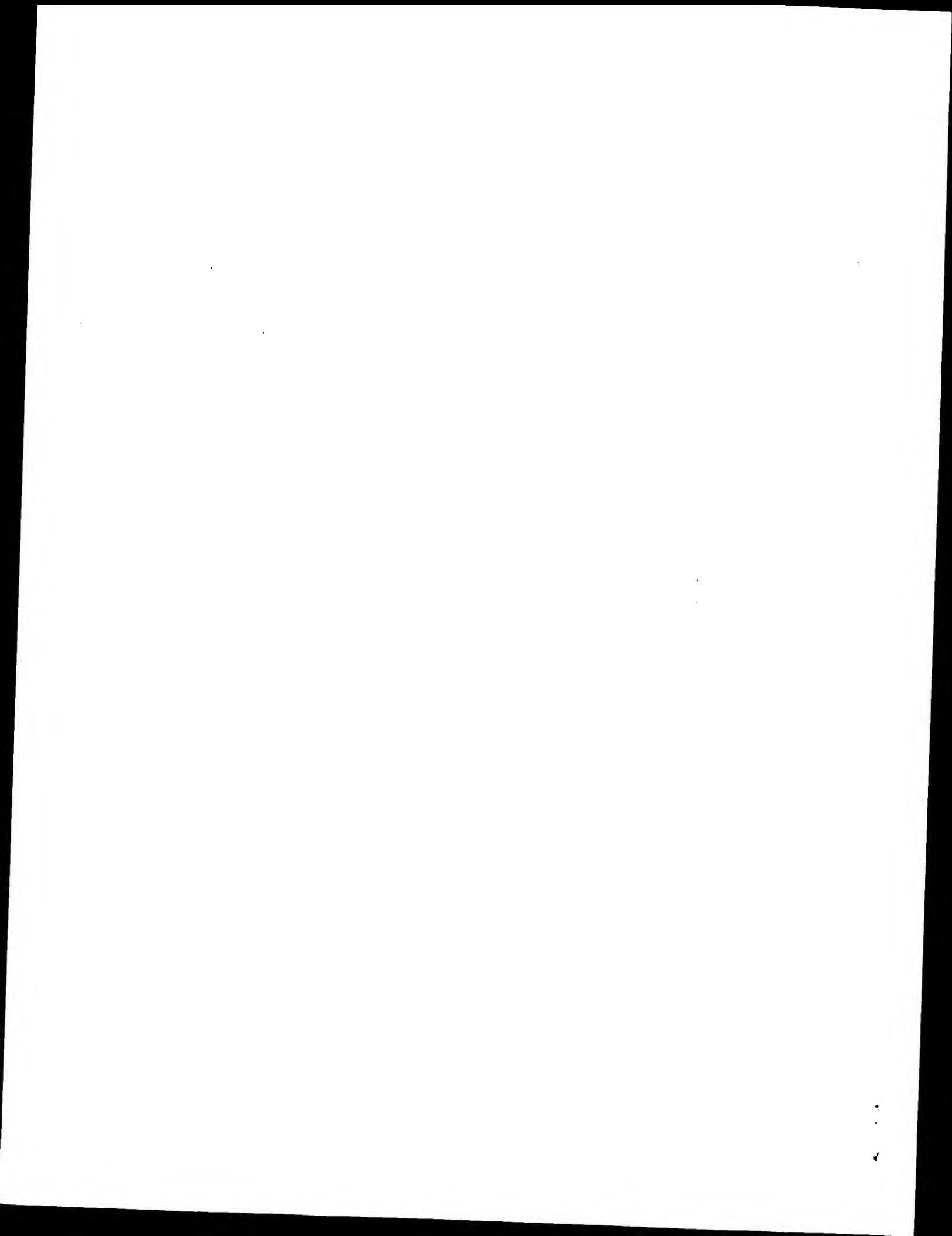
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 EAEXIXXEEIVPNXVEQK 21

DB 18 ESERIESSRELVDNEEEK 36

Search completed: February 11, 2003, 18:19:40

Job time : 26.5979 secs



Wed Feb 12 11:35:21 2003

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:06:55 ; Search time 36.3402 Seconds
(without alignments)
91.669 Million cell updates/sec

Title: US-09-380-738a-2

Perfect score: 105

Sequence: 1 RELEENVPGEIVLXXXXESIR 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	96.2	25	AA28425	Anticariogenic pho
2	101	96.2	25	AA28425	Casein phosphopept
3	101	96.2	25	AA28425	Sequence of casein
4	101	96.2	25	AA28425	Bos beta-casein X-
5	101	96.2	209	AA28425	Methyl or ethyl es
6	101	96.2	209	AA28425	Bovine beta casein
7	101	96.2	209	AA28425	Bovine beta-casein
8	96	91.4	24	AA28425	Phosphopeptide #2
9	96	91.4	24	AA28425	Anticariogenic pho
10	96	91.4	24	AA28425	Anticariogenic pho

11	96	91.4	24	14	AA28425	Casein phosphopept
12	96	91.4	24	14	AA28425	phosphopeptide 2.
13	96	91.4	24	14	AA28425	Sequence of casein
14	96	91.4	24	15	AA28425	Sodium caseinate t
15	96	91.4	24	16	AA28425	Phosphopeptide 12.
16	96	91.4	24	21	AA28425	Casein phosphopept
17	96	91.4	28	17	AA28425	Modified Casein se
18	96	91.4	209	19	AA28425	Beta-casein. Capr
19	96	91.4	222	17	AA28425	Disaccharide decem
20	45	42.9	213	21	AA28425	Human beta-casein.
21	45	42.9	213	21	AA28425	Solanum tuberosum
22	43	41.0	974	19	AA28425	Venezuelan equine
23	43	41.0	2492	23	AA28425	H. pylori GHPO 363
24	42.5	40.5	140	19	AA28425	Beta-casein. Homo
25	42	40.0	225	14	AA28425	Novel human diagno
26	42	40.0	262	22	AA28425	Salmonella typhi c
27	42	40.0	304	22	AA28425	Drosophila melanog
28	42	40.0	522	22	AA28425	Arabidopsis thalia
29	41	39.0	113	21	AA28425	Arabidopsis thalia
30	41	39.0	141	21	AA28425	Arabidopsis thalia
31	41	39.0	178	21	AA28425	A. thaliana FLF-11
32	41	39.0	196	21	AA28425	Amino acid sequenc
33	41	39.0	196	22	AA28425	Propionibacterium
34	41	39.0	221	22	AA28425	Arabidopsis thalia
35	41	39.0	228	21	AA28425	Arabidopsis thalia
36	41	39.0	265	21	AA28425	C. glutamicum prote
37	41	39.0	275	22	AA28425	Staphylococcus epi
38	41	39.0	1037	23	AA28425	Novel human secret
39	40.5	38.6	77	22	AA28425	Sequence encoded b
40	40	38.1	225	14	AA28425	Human beta-casein.
41	40	38.1	225	19	AA28425	pre-pro-VGL. Xeno
42	40	38.1	358	15	AA28425	Arabidopsis thalia
43	40	38.1	520	21	AA28425	Herbicideally activ
44	40	38.1	520	23	AA28425	Putative P. abyss
45	40	38.1	613	22	AA28425	

ALIGNMENTS

RESULT 1
AA28425
ID AA28425 standard; peptide; 25 AA.

XX AA28425;

XX 19-MAR-1993 (first entry)

XX Anticariogenic phosphopeptide.

DE Casein; metal ion; dietetic; purification; growth medium; dietary supplement; fertiliser.

XX Key Location/Qualifiers

FT Modified-site 15

FT /note= "post-translationally phosphorylated serine"

FT Modified-site 17

FT /note= "post-translationally phosphorylated serine"

FT Modified-site 18

FT /note= "post-translationally phosphorylated serine"

FT Modified-site 19

FT /note= "post-translationally phosphorylated serine"

XX WO9218526-A.

XX 29-OCT-1992.

XX 16-APR-1992; 92WO-AU00175.

XX 19-APR-1991; 91AU-0005706.

XX (UYME) UNIV MELBOURNE.

XX (VICT-) VICTORIAN DAIRY IND AUTHORITY.

XX PI Reynolds EC;
 XX DR WPI; 1992-382039/46.
 XX PT Selected phospho-peptide(s) prodn. having anticariogenic
 XX PT activities etc. - comprises digesting soluble monovalent cation
 XX PT salt of casein in soln.; introducing di- or trivalent metal ion
 XX PT and filtering through filter having mol. wt. exclusion limit
 XX PS Claim 11; Page 9; 18pp; English.
 XX CC The peptide may be prep. by completely digesting casein in soln.
 XX CC with a proteolytic enzyme, adding mineral acid to the soln. to
 XX CC adjust the pH to 4.7, removing any precipitate, adding calcium
 XX CC chloride to cause aggregation of the peptides in soln. and
 XX CC separating the aggregated phosphopeptides. This method allows prodn.
 XX CC of the phosphopeptide by industrial methods. The phosphopeptide
 XX CC has anticariogenic activity and may be used as a dietetic. The
 XX CC peptide may be used in a microbiological growth medium, as a dietary
 XX CC supplement or as a fertiliser.
 XX CC See also AAR28426-33.
 XX SQ Sequence 25 AA;
 Query Match 96.2%; Score 101; DB 13; Length 25;
 Best Local Similarity 84.0%; Pred. No. 2.4e-10;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 RELEELNVGPEIVEXLXXXEESITR 25
 DB 1 RELEELNVGPEIVESLSSEESITR 25
 RESULT 2
 AAR32927
 ID AAR32927 standard; Protein; 25 AA.
 AC AAR32927;
 XX DT 02-JUL-1993 (first entry)
 XX DE Casein phosphopeptide #1.
 XX KW Casein; phosphopeptide; dental calculus; salts; alkaline metal;
 XX KW alkaline earth metal; Zn/phosphopeptide complex; aggregate;
 XX KW anti-carries; anti-gingivitis.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 XX FT Modified-site 15
 XX FT Modified-site 17 /label= Phosphoserine
 XX FT Modified-site 18 /label= Phosphoserine
 XX FT Modified-site 19 /label= Phosphoserine
 XX FT Modified-site 19 /label= Phosphoserine
 XX PN WO9303707-A.
 XX PD 04-MAR-1993.
 XX PF 21-AUG-1992; 92WO-AU000441.
 XX PR 22-AUG-1991; 91US-0748344.
 XX PA (UYME) UNIV MELBOURNE.
 XX PA (VICT-) VICTORIAN DAIRY IND AUTHORITY.
 XX PI Reynolds EC;

DR WPI; 1993-093685/11.
 XX PT Controlling dental calculus by treating teeth with oral compsns.
 XX PT - which contains phospho-peptide(s) having 5-40 amino-acyl
 XX PT residues
 XX PS Claim 3; Page 15; 23pp; English.
 XX CC The sequences given in AAR32927-35 are casein phosphopeptides which can
 XX CC be used to inhibit dental calculus. These peptides are pref. in the
 XX CC form of salts selected from alkaline metal, alkaline earth metal salts
 XX CC such as Na, Ca, Zn, Cu, Al, K, Sr, Mg and Ni salts. These peptides
 XX CC are pref. present as a Zn/phosphopeptide complex or aggregate. These
 XX CC peptides have anti-calculus potential, and are anti-carries and anti-
 XX CC gingivitis agents.
 XX SQ Sequence 25 AA;
 Query Match 96.2%; Score 101; DB 14; Length 25;
 Best Local Similarity 84.0%; Pred. No. 2.4e-10;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 RELEELNVGPEIVEXLXXXEESITR 25
 DB 1 RELEELNVGPEIVESLSSEESITR 25
 RESULT 3
 AAR47814
 ID AAR47814 standard; peptide; 25 AA.
 AC AAR47814;
 XX DT 21-JUL-1994 (first entry)
 XX DE Sequence of casein phosphopeptide (CPP).
 XX KW Casein phosphopeptide; dental hypersensitivity; therapy; tooth.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 XX FT Modified-site 15 /label= Ser(P)
 XX FT /note= *see also residues 17,18,19"
 XX PN WO9400146-A.
 XX PD 06-JAN-1994.
 XX PF 29-JUN-1993; 93WO-AU00319.
 XX PR 29-JUN-1992; 92AU-0003221.
 XX PA (UYME) UNIV MELBOURNE.
 XX PA (VICT-) VICTORIAN DAIRY IND AUTHORITY.
 XX PI Reynolds EC;
 XX DR WPI; 1994-025888/03.
 XX PT Treatment of dental hypersensitivity - using casein, component
 XX PT of casein, phospho-protein or phospho-peptide or their salts
 XX PS Disclosure; Page 11; 23pp; English.
 XX CC Ser(P) - post-translationally phosphorylated serine. A mixture of
 XX CC casein phosphopeptides (CPP) and/or their salts may be used in a
 XX CC method for treating dental hypersensitivity. Pref. those CPPs
 XX CC contg. the sequence -Ser(P)-Ser(P)-Ser(P)-predominate. The CPPs
 XX CC can be extracted from a casein digest.
 XX SQ Sequence 25 AA;

Wed Feb 12 11:35:21 2003

CC cats and dogs as well as zoo animals. The present sequence represents a
 CC phosphopeptide component of a specifically claimed complex.
 XX
 SQ Sequence 25 AA; Query Match 96.2%; Score 101; DB 15; Length 25;
 Best Local Similarity 84.0%; Pred. No. 2.4e-10; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 RELEELNVPGEIVEXLXXEESITR 25
 Db 1 RELEELNVPGEIVESLSSESITR 25
 RESULT 4
 AAW66600
 ID AAW66600 standard; peptide; 25 AA.
 XX
 AC AAW66600;
 DT 27-NOV-1998 (first entry)
 XX Bos beta-casein X-4P (fl-25) phosphopeptide.
 DE casein; calcium phosphate complex; amorphous calcium phosphate; ACP;
 KW phosphopeptide; delivery vehicle; calcium fluoride; calcium deficiency;
 KW osteoporosis; osteomalacia; tooth; bone disease.
 XX Synthetic.
 OS Bos taurus.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 15 /note= "Ser(P)"
 FT Modified-site 17 /note= "Ser(P)"
 FT Modified-site 18 /note= "Ser(P)"
 FT Modified-site 19 /note= "Ser(P)"
 FT Modified-site 19 /note= "Ser(P)"
 FT
 FT WO9840406-A1.
 XX
 XX 17-SEP-1998.
 XX 13-MAR-1998; 98WO-AU00160.
 XX 13-MAR-1997; 97AU-0005662.
 XX (UYME) UNIV MELBOURNE.
 XX (VICT-) VICTORIAN DAIRY IND AUTHORITY.
 XX Reynolds EC;
 XX WPI; 1998-520803/44.
 XX
 XX Stable calcium phosphate complex including phospho:peptide
 PT stabilised amorphous calcium phosphate - useful for treatment of
 PT dental caries, calcium malabsorption and bone diseases such as
 PT osteoporosis and osteomalacia.
 XX
 XX Claim 4; Page 35; 43pp; English.
 XX
 XX The invention relates to a stable calcium phosphate complex including
 CC phosphopeptide stabilised amorphous calcium phosphate (ACP) or its
 CC derivative, where the phosphopeptide includes the amino acid sequence:
 CC Ser(P)-Ser(P)-Glu-Glu. The amorphous phases stabilised by the
 CC phosphopeptides are a delivery vehicle for co-localisation of Ca,
 CC P and phosphate at the tooth surface in a slow-release amorphous form
 CC producing superior anticaries efficacy over prior art. The amorphous
 CC phases stabilised by the phosphopeptides are also useful as dietary
 CC supplements to increase calcium bioavailability and to help prevent
 CC diseases associated with calcium deficiencies. They are particularly
 CC useful for treatment or prevention of dental caries, calcium
 CC malabsorption and bone diseases such as osteoporosis and osteomalacia.
 CC The compositions are useful in humans and in veterinary medicine in
 CC domestic animals such as cattle, sheep, horses and companion animals e.g.

```

FT FT /note= "pepsin cleavage site in native protein"
FT FT 127..128
FT FT /note= "pepsin cleavage site in native protein"
FT FT 141..142
FT FT /note= "pepsin cleavage site in native protein"
FT FT 142..143
FT FT /note= "pepsin cleavage site in native protein and
FT FT in methyl and ethyl esters of beta-casein"
FT FT 156..157
FT FT /note= "newly identified pepsin cleavage site in
FT FT ethyl ester of beta-casein"
FT FT 162..163
FT FT /note= "newly identified pepsin cleavage site in
FT FT ethyl ester of beta-casein"
FT FT 163..164
FT FT /note= "pepsin cleavage site in native protein and
FT FT in methyl and ethyl esters of beta-casein"
FT FT 164..165
FT FT /note= "pepsin cleavage site in native protein and
FT FT in ethyl ester of beta-casein"
FT FT 188..189
FT FT /note= "pepsin cleavage site in native protein and
FT FT in methyl ester of beta-casein"
FT FT 189..190
FT FT /note= "pepsin cleavage site in native protein and
FT FT in methyl and ethyl esters of beta-casein"
FT FT 190..191
FT FT /note= "pepsin cleavage site in native protein and
FT FT in ethyl ester of beta-casein"
FT FT 191..192
FT FT /note= "pepsin cleavage site in native protein and
FT FT in methyl and ethyl esters of beta-casein"
FT FT 192..193
FT FT /note= "pepsin cleavage site in native protein and
FT FT in methyl and ethyl esters of beta-casein"
FT FT 198..199
FT FT /note= "newly identified pepsin cleavage site in
FT FT methyl and ethyl esters of beta-casein"
FT FT 207..208
FT FT /note= "newly identified pepsin cleavage site in
FT FT methyl ester of beta-casein"
FT FT 2..25
FT FT /label= A
FT FT /note= "tryptic peptide from native protein"
FT FT 26..28
FT FT /label= B
FT FT /note= "tryptic peptide from native protein"
FT FT 29..32
FT FT /label= C
FT FT /note= "tryptic peptide from native protein"
FT FT 33..48
FT FT /label= D
FT FT /note= "tryptic peptide from native protein"
FT FT 49..97
FT FT /label= E
FT FT /note= "tryptic peptide from native protein"
FT FT 100..105
FT FT /label= F
FT FT /note= "tryptic peptide from native protein"
FT FT 106..107
FT FT /label= G
FT FT /note= "tryptic peptide from native protein"
FT FT 108..113
FT FT /label= H
FT FT /note= "tryptic peptide from native protein"
FT FT 114..169
FT FT /label= I
FT FT /note= "tryptic peptide from native protein"
FT FT 170..176
FT FT /label= J
FT FT /note= "tryptic peptide from native protein"
FT FT 177..183
FT FT /label= K

```

```

FT Peptide /note= "tryptic peptide from native protein"
FT 184..202 /label= L
FT /note= "tryptic peptide from native protein"
FT 203..209 /label= N
FT /note= "tryptic peptide from native protein"
FT 15 /note= "phosphorylated"
FT 17 /note= "phosphorylated"
FT 18 /note= "phosphorylated"
FT 19 /note= "phosphorylated"
FT 35 /note= "phosphorylated"
FT /note= "phosphorylated"
FT WO9517518-A1.
FT PN
FT XX
FT PD 29-JUN-1995.
FT XX
FT PF 20-DEC-1994; 94WO-FR01500.
FT XX
FT PR 23-DEC-1993; 93FR-0015764.
FT XX
FT PA (INRG ) INST NAT RECH AGRONOMIQUE.
FT XX
FT PI Briand L, Chobert J, Haertle T;
FT DR WPI; 1995-240679/31.
FT XX
FT PT New esterified amino acids, peptide(s) and their mixts. - prepd. by
FT esterification of protein then enzymatic hydrolysis, useful as
FT ingredients and additives in foods, pharmaceuticals and cosmetics
FT XX
FT PS Claim 7; Fig 7 and 18; 47pp; French.
FT XX
FT CC The native form of bovine beta-casein A1 contains various pepsin
FT cleavage sites. Esterification of the protein with methanol or ethanol
FT results in a form of beta-casein contg. additional, non-conventional
FT pepsin cleavage sites (see Features Table). Esterified peptides and
FT amino acids (and their mixtures) resulting from hydrolysis of an
FT esterified protein (pref. beta-lactoglobulin or beta-casein) are
FT claimed. The hydrolysis products are useful as ingredients,
FT additives or active agents in foods, pharmaceuticals and cosmetics.
FT XX
FT SQ Sequence 209 AA;
FT
FT Query Match 96.2%; Score 101; DB 16; Length 209;
FT Best Local Similarity 84.0%; Pred. No. 3e-09;
FT Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
FT
FT QY 1 RELEELNVPGEIVEXLXXEESITR 25
FT | | | | | | | | | | | | | | | | | | | |
FT Db 1 RELEELNVPGEIVESLSSESITR 25
FT
FT RESULT 6
FT AAR95609
FT ID AAR95609 standard; protein; 209 AA.
FT XX
FT AC AAR95609;
FT XX
FT DT 26-NOV-1996 (first entry)
FT XX
FT DE Bovine beta casein A1 variant.
FT XX
FT KW Milk; beta casein; diabetogenic; diabetes; cow; milk products;
FT butter; cheese; cream.
FT XX
FT OS Bos taurus.
FT XX

```

us-09-380-738a-2.rag

Wed Feb. 12 11:35:21 2003

WPI; 2001-112363/12.

Prophylactic dietary supplement for reducing incidence of cardio, cerebro vascular diseases and diabetes in a population, contains milk or its products fortified with betaine, cobalamin, folic acid or pyridoxine

Claim 9; Fig 3; 32pp; English.

The invention relates to a dietary supplement which, when consumed, is capable of reducing plasma levels of homocyst(e)ine (thcy). thcy is a major risk indicator of heart disease and vascular disease in general in humans. Vascular wall health is also seriously compromised in patients with clinical or unrecognised diabetes, with thcy being a strong risk factor for mortality in type II diabetic patients. Deficiencies in folic acid, pyridoxine and cobalamin lead to higher thcy levels, and folic acid deficiency is known to be involved in vascular disease, as well as causing neural tube defects in early embryonic development. The dietary supplement of the invention comprises milk or a milk product, fortified by the addition of betaine, cobalamin, folic acid, pyridoxine or their analogues. In addition, the beta-casein component of the milk is substantially the A2 variant. Beta-casein types A1 and B, consumption of which are correlated with the incidence of type I diabetes, are substantially excluded from the supplement. The dietary supplement is useful for reducing the incidence of vascular disease, including and peripheral vascular disease and blood vessel wall degeneration and particularly cardiovascular disease and cerebrovascular disease, and is also useful for reducing the incidence of type I and II diabetes. It additionally provides a sufficient daily dose of folic acid to prevent neural tube defects in foetuses. The supplement provides health improvements to a human population without the administration of medication. The present sequence represents bovine beta-casein type A2.

Sequence 209 AA;

Query Match 96.2%; Score 101; DB 22; Length 209;

Best Local Similarity 84.0%; Pred. No. 3e-09;

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RELEELNVPGEIVEXLXXXESITR 25

DB 1 RELEELNVPGEIVESLSSESITR 25

RESULT 8

.. AAF71321

ID AAF71321 standard; peptide; 24 AA.

XX AC AAF71321;

XX DT 07-MAY-1991 (first entry)

XX DE Phosphopeptide 2.

XX KW Caries; gingivitis; periodontal disease; osteoporosis; osteomalacia.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 14 /label= phosphoserine

FT Modified-site 16 /label= phosphoserine

FT Modified-site 17 /label= phosphoserine

FT Modified-site 18 /label= phosphoserine

FT Modified-site 18 /label= phosphoserine

XX WO8707616-A.

PN 17-DEC-1987.

XX 12-JUN-1987;

PF 87WO-AU00172.

FH Key Location/Qualifiers

FT Region 63..68

XX /label= Diabetogenic hexapeptide.

PN WO9614577-A1.

XX 17-MAY-1996.

PD 03-NOV-1995; 95WO-NZ00114.

PE 04-NOV-1994; 94NZ-0264862.

PR (NACH-) NAT CHILD HEALTH RES FOUND.

PA (NZDA-) NEW ZEALAND DAIRY BOARD.

XX Elliott RB, Hill JP;

XX WPI; 1996-251885/25.

DR Selecting non-diabetogenic milk and milk prods. - by testing milk or

PT cows for the presence of non-diabetogenic variants of beta-casein

PS Disclosure; Figure 2; 28pp; English.

XX A method for selecting milk for feeding to diabetes susceptible

CC individuals comprises testing milk from identified cows for the

CC presence of variants of beta casein and selecting those cows whose

CC milk contains non-diabetogenic variants and milking these cows

CC separately. The milk and milk products obtained can reduce the risk

CC of susceptible individuals contracting type-1 diabetes.

XX Sequence 209 AA;

QY 1 RELEELNVPGEIVEXLXXXESITR 25

DB 1 RELEELNVPGEIVESLSSESITR 25

RESULT 7

AAB60348

ID AAB60348 standard; protein; 209 AA.

XX AC AAB60348;

XX DT 06-APR-2001 (first entry)

XX DE Bovine beta-casein type A2.

XX KW Bovine; beta-casein; type A2; milk; dairy product; dietary supplement;

XX KW plasma homocysteine; homocystine; thcy; folic acid; pyridoxine; betaine;

XX KW cobalamin; vascular disease; diabetes; peripheral vascular disease;

XX KW blood vessel wall degeneration; cardiovascular disease;

XX KW cerebrovascular disease; neural tube defect; prophylaxis.

XX OS Bos taurus.

XX WO200100047-A1.

XX 04-JAN-2001.

XX 29-JUN-2000; 2000WO-NZ00116.

XX 29-JUN-1999; 99NZ-0336505.

PR 18-APR-2000; 2000NZ-0504057.

XX (NZMI-) NEW ZEALAND MILK INST LTD.

PA Elliott RB, Laugesen BM;

XX 12-JUN-1987;

PF 87WO-AU00172.

XX PR 12-JUN-1986; 86AU-0006385.
 XX PA (VICT-) VICTORIA DAIRY INDUSTRY AUTHORITY.
 XX PA (UYME-) UNIVERSITY OF MELBOURNE.
 XX PA (REYN/) EC REYNOLDS.
 XX PI Reynolds EC;
 XX DR WPI; 1987-362707/51.
 XX PT New phosphopeptides contg. defined amino acid sequence - useful in
 XX PT treatment of dental, rarefying bone diseases and disease relating to
 XX PT malabsorption of minerals.
 XX PS Claim 5; Page 17; 22pp; English.
 XX CC The phosphopeptide is used in compsns. at a conc. of 0.01-5 wt%.
 XX CC See also AAP71320-P71324.
 XX SQ Sequence 24 AA;

Query Match 91.4%; Score 96; DB 8; Length 24;
 Best Local Similarity 83.3%; Pred. No. 1.6e-09;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 EELNVPGEIVEXLXXEESITR 25
 DB 1 EELNVPGEIVESLSSESITR 24

RESULT 9
 AAR14448
 ID AAR14448 standard; Protein; 24 AA.
 AC AAR14448;
 DT 10-JAN-1992 (first entry)
 DE Phosphopeptide #2 for increasing mineral absorption.
 KW caries; gingivitis; periodontal disease; osteoporosis; osteomalacia;
 KW anaemia.
 OS Synthetic.

Key	Location/Qualifiers
Modified-site 14	/label= OTHER
Modified-site 16	/note= "phosphoserine (Pse)"
Modified-site 17	/label= OTHER
Modified-site 18	/note= "Pse"
Modified-site 19	/label= OTHER
Modified-site 20	/note= "Pse"

US5015628-A.
 14-MAY-1991.
 03-AUG-1990; 90US-0563798.
 03-AUG-1990; 90US-0563798.

(UYME-) UNIV OF MELBOURNE.
 (VICT-) VICTORIAN DAIRY INDUSTRY.
 Reynolds EC;

XX

XX WPI; 1991-316875/43.
 XX Novel phosphopeptides - useful for treating dental diseases,
 XX PT rarefying diseases or diseases relating to malabsorption of minerals
 XX PS Claim 1; Column 11; 8pp; English.
 XX CC This is an example of a highly generic formula for a phosphopeptide
 XX CC of length 5-13 amino acids. The peptides can be made synthetically
 XX CC (e.g. chemical synthesis or genetic engineering) or they can be
 XX CC extracted from cereals, nuts or vegetables or by fractionating a
 XX CC digest of casein, alpha-s-casein, beta-casein or a salt of it.
 XX CC Compositions comprising the peptide may take the form of foodstuff
 XX CC or confectionery, dentifrices, mouthwashes and preparations for
 XX CC topical application to teeth or gingival tissue. The peptides
 XX CC significantly increase absorption of calcium, phosphate and iron in
 XX CC the gut. See AAR14447-R14451.
 XX SQ Sequence 24 AA;

Query Match 91.4%; Score 96; DB 12; Length 24;
 Best Local Similarity 83.3%; Pred. No. 1.6e-09;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 EELNVPGEIVEXLXXEESITR 25
 DB 1 EELNVPGEIVESLSSESITR 24

RESULT 10
 AAR28427
 ID AAR28427 standard; peptide; 24 AA.
 AC AAR28427;
 DT 19-MAR-1993 (first entry)
 DE Anticariogenic phosphopeptide.
 KW Casein; metal ion; dietetic; purificaiton; growth medium; dietary
 KW supplement; fertiliser.
 XX Key Location/Qualifiers

Key	Location/Qualifiers
Modified-site 14	/note= "post-translationally phosphorylated serine"
Modified-site 16	/note= "post-translationally phosphorylated serine"
Modified-site 17	/note= "post-translationally phosphorylated serine"
Modified-site 18	/note= "post-translationally phosphorylated serine"
Modified-site 19	/note= "post-translationally phosphorylated serine"

WO9218526-A.
 29-OCT-1992.
 16-APR-1992; 92WO-AU00175.
 19-APR-1991; 91AU-0005706.
 (UYME) UNIV MELBOURNE.
 (VICT-) VICTORIAN DAIRY IND AUTHORITY.
 Reynolds EC;
 WPI; 1992-382039/46.

Selected phospho-peptide(s) prodn. having anticariogenic
 activities etc. - comprises digesting soluble monovalent cation
 salt of casein in soln., introducing di- or trivalent metal ion
 and filtering through filter having mol. wt. exclusion limit
 Claim 11; Page 11; 18pp; English.

us-09-380-738a-2.rag

Wed Feb. 12 11:35:21 2003

XX The peptide may be prep. by completely digesting casein in soln.
 CC with a proteolytic enzyme, adding mineral acid to the soln. to
 CC adjust the pH to 4.7, removing any precipitate, adding calcium
 CC chloride to cause aggregation of the peptides in soln. and
 CC separating the aggregated phosphopeptides. This method allows prodn.
 CC of the phosphopeptide by industrial methods. The phosphopeptide
 CC has anticariogenic activity and may be used as a dietetic. The
 CC peptide may be used in a microbiological growth medium, as a dietary
 CC supplement or as a fertilizer.
 CC See also AAR28425-33.
 XX
 XX SQ Sequence 24 AA; 91.4%; Score 96; DB 13; Length 24;
 Query Match 83.3%; Pred. No. 1.6e-09;
 Best Local Similarity 83.3%; Mismatches 4; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ELEELNVPGEIVEXLXXXESITR 25
 DB 1 ELEELNVPGEIVESLSSESITR 24
 RESULT 11
 AAR32929
 ID AAR32929 standard; Protein; 24 AA.
 XX
 XX AC AAR32929;
 XX
 XX 02-JUL-1993 (first entry)
 XX
 XX Casein phosphopeptide #3.
 XX Casein; phosphopeptide; dental calculus; salts; alkaline metal;
 XX alkaline earth metal; Zn/phosphopeptide complex; aggregate;
 XX anti-carries; anti-gingivitis.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Modified-site 14 /label= Phosphoserine
 FT Modified-site 16 /label= Phosphoserine
 FT Modified-site 17 /label= Phosphoserine
 FT Modified-site 18 /label= Phosphoserine
 FT Modified-site 18 /label= Phosphoserine
 XX WO9303707-A.
 XX
 XX 04-MAR-1993.
 XX
 XX 21-AUG-1992; 92WO-AU00441.
 XX
 XX 22-AUG-1991; 91US-0748344.
 XX (UYME) UNIV MELBOURNE.
 XX (VICT-) VICTORIAN DAIRY IND AUTHORITY.
 XX Reynolds EC;
 XX WPT; 1993-093685/11.
 XX Controlling dental calculus by treating teeth with oral compans.
 XX - which contains phospho-peptide(s) having 5-40 amino-acyl
 XX residues
 XX Claim 3; Page 16; 23pp; English.
 XX The sequences given in AAR32927-35 are casein phosphopeptides which can
 CC be used to inhibit dental calculus. These peptides are pref. in the
 CC form of salts selected from alkaline metal, alkaline earth metal salts

CC such as Na, Ca, Zn, Cu, Al, K, Sr, Mg and Ni salts. These peptides
 CC are pref. present as a Zn/phosphopeptide complex or aggregate. these
 CC peptides have anti-calculus potential, and are anti-carries and anti-
 CC gingivitis agents.
 XX
 XX SQ Sequence 24 AA; 91.4%; Score 96; DB 14; Length 24;
 Query Match 83.3%; Pred. No. 1.6e-09;
 Best Local Similarity 83.3%; Mismatches 4; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ELEELNVPGEIVEXLXXXESITR 25
 DB 1 ELEELNVPGEIVESLSSESITR 24
 RESULT 12
 AAR31238
 ID AAR31238 standard; peptide; 24 AA.
 XX
 XX AC AAR31238;
 XX
 XX 18-MAY-1993 (first entry)
 XX
 XX Phosphopeptide 2.
 XX Phosphopeptide; active agent; oral; composition; anionic; polymeric;
 KW stabiliser; carboxylate; polymer; sulfonate; destabilisation;
 KW fluoride; carries; gingivitis.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Modified-site 14 /label= Phosphoserine
 FT Modified-site 16 /label= Phosphoserine
 FT Modified-site 17 /label= Phosphoserine
 FT Modified-site 18 /label= Phosphoserine
 FT Modified-site 18 /label= Phosphoserine
 XX EP523776-A.
 XX
 XX 20-JAN-1993.
 XX
 XX 03-JUL-1992; 92EP-0202023.
 XX
 XX 17-JUL-1991; 91US-0731592.
 XX (UNIL) UNILEVER NV.
 XX (UNIL) UNILEVER PLC.
 XX Burger AR, Elliott DL, Schick LA;
 XX WPI; 1993-019802/03.
 XX Oral compans. contg. a phospho-peptide - with addn. of an anionic
 XX polymeric stabiliser to inhibit destabilisation in the oral
 XX environment
 XX Claim 6; Page 14; 18pp; English.
 XX The sequences given in AAR31237-42 represent phosphopeptides which
 CC were used as the active agents in an oral composition. these
 CC peptides were stabilised by an anionic polymeric stabiliser. The
 CC anionic polymers were chosen from a group consisting of carboxylate
 CC polymers, sulfonate polymers, polymers having both a carboxylate and
 CC a sulfonate moiety, and other mixtures. The anionic polymeric
 CC stabiliser inhibits destabilisation of the phosphopeptide in the oral
 CC environment. These oral compositions, pref. containing a fluoride
 CC source may be used for inhibiting carries and gingivitis.

1

us-09-380-738a-2.rag

Wed Feb 12 11:35:21 2003

sugar; sorbitol; mannitol; xylitol; lactitol; cellobiitol; confectionary;
caries; gingivitis; calcium; remineralisation.

Bos taurus.

Key Location/Qualifiers
Modified-site 14 /label= OTHER
/note= "Phosphoserine"
Modified-site 16 /label= OTHER
/note= "Phosphoserine"
Modified-site 17 /label= OTHER
/note= "Phosphoserine"
Modified-site 18 /label= OTHER
/note= "Phosphoserine"

EP629393-A.

21-DEC-1994.

07-JUN-1994; 94EP-0304083.

16-JUN-1993; 93US-0078706.

(ICIL) ICI AMERICAS INC.

Duross JW;

WPI; 1995-044845/07.

Crystalline poly:ol compositions for use in confectionery and
pharmaceuticals - contain a phospho:peptide uniformly
distributed within the crystal matrix and can improve oral
hygiene

Disclosure; column 4; 8pp; English.

Peptides (AAR68936-40) are phosphopeptides resulting from the tryptic
digestion of sodium caseinate. The peptide shown here is derived from
the TCPK-tryptic digest of beta-caseinate. The peptides contain the
amino acids phosphoserine, phosphothreonine, phosphotyrosine or
phosphohistidine, and include in their sequence aspartate and glutamate.
The phosphopeptides form part of a novel composition containing the
peptide dispersed uniformly in a crystal mix of sugars chosen from
sorbitol, mannitol, xylitol, lactitol, cellobiitol or mixtures of
sorbitol/mannitol or sorbitol/xylitol. The compositions can be used
instead of conventional sugars to manufacture confectionary products.
The phosphopeptides are known to inhibit caries and gingivitis and can
act as a source of calcium ions to promote/enhance remineralisation.

Sequence 24 AA;

Query Match 91.4%; Score 96; DB 15; Length 24;
Best Local Similarity 83.3%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 EEEENVPGEIVEIXXXEESITR 25
Db 1 EEEENVPGEIVESLSSEISITR 24

Search completed: February 11, 2003, 18:16:54
Job time : 37.3402 secs

us-09-380-738a-2.ra1

Wed Feb -12 11:35:21 2003

Sequence 3, Appli
Sequence 5016, Ap
Sequence 1, Appli
Sequence 3, Appli
Sequence 2, Appli
Sequence 22, Appli
Sequence 21, Appli
Sequence 40, Appli
Sequence 2, Appli
Sequence 5, Appli
Sequence 2, Appli
Sequence 7, Appli
Sequence 5, Appli
Sequence 2, Appli
Sequence 53, Appli
Sequence 20, Appli
Sequence 8, Appli

28 39 37.1 2523 4 US-08-899-232-3
29 38 36.2 270 4 US-09-134-001C-5016
30 38 36.2 271 4 US-09-414-436-1
31 37 35.2 16 4 US-08-341-555-3
32 37 35.2 51 2 US-08-312-662-2
33 37 35.2 54 4 US-08-341-555-22
34 37 35.2 99 4 US-08-492-411A-40
35 37 35.2 196 1 US-08-387-845-2
36 37 35.2 196 2 US-08-999-811-5
37 37 35.2 196 2 US-08-778-275-2
38 37 35.2 196 2 US-08-824-966-7
39 37 35.2 196 3 US-09-042-105-5
40 37 35.2 196 3 US-08-867-352-2
41 37 35.2 196 4 US-08-795-430-53
42 37 35.2 196 4 US-09-355-700-53
43 37 35.2 196 5 PCT-US91-02766-20
44 37 35.2 196 5 PCT-US96-09001-8
45 37 35.2 196 5

ALIGNMENTS

RESULT 1
US-07-748-344B-1
; Sequence 1, Application US/07748344B
; Patent No. 5227154
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; TITLE OF INVENTION: PHOSPHOPEPTIDES FOR THE
; TREATMENT OF DENTAL CALCULUS
; NUMBER OF INVENTIONS: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDRUS, SCEALES, STARKE & SAWALL
; STREET: 100 EAST WISCONSIN AVE., SUITE 1100
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07748, 344B
; FILING DATE: 19910822
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: SARA, CHARLES S
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: C.8493-87
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 255-2182
; TELEFAX: (608) 255-2182
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 15
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION: Phosphoserine
; NAME/KEY: Phosphoserine

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:14:31 : Search time 12.3711 Seconds
(without alignments)
59.459 Million cell updates/sec

Title: US-09-380-738A-2

Perfect score: 105

Sequence: 1 RELELNVPGEIVELXXXXESITR 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	96.2	25	1	US-07-748-344B-1
2	101	96.2	25	2	US-08-621-564B-2
3	101	96.2	25	2	US-08-954-985A-1
4	101	96.2	25	4	US-08-137-086-1
5	101	96.2	209	4	US-09-269-220-2
6	101	96.2	209	4	US-08-836-778-2
7	96	91.4	24	1	US-07-748-344B-3
8	96	91.4	24	2	US-08-954-985A-3
9	96	91.4	24	4	US-08-137-086-3
10	96	91.4	28	2	US-08-621-564B-1
11	96	91.4	209	4	US-09-269-220-1
12	76	72.4	222	2	US-08-391-743A-2
13	76	72.4	222	4	US-09-143-155-2
14	45	42.9	213	4	US-09-131-028A-2
15	45	42.9	213	4	US-09-131-028A-12
16	43	41.0	974	2	US-08-868-786-6
17	41	39.0	1037	4	US-09-134-001C-4794
18	40	38.1	360	2	US-08-459-346-13
19	40	38.1	360	2	US-08-411-607A-4
20	40	38.1	360	3	US-08-889-419-13
21	40	38.1	360	4	US-08-402-542-13
22	40	38.1	360	5	PCT-US93-07189-13
23	39	37.1	210	1	US-08-078-090-2
24	39	37.1	519	4	US-09-453-702B-265
25	39	37.1	1064	1	US-08-537-210A-3
26	39	37.1	1064	1	US-09-113-825-3
27	39	37.1	2523	1	US-08-185-432-18

us-09-380-738a-2.rai

Wed, Feb-12 11:35:21 2003

```

APPLICANT: REYNOLDS, ERIC CHARLES
TITLE OF INVENTION: PRODUCTION OF PHOSPHOPEPTIDES FROM CASEIN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: ANDRUS, SCALES, STARKE & SAWALL
STREET: 100 EAST WISCONSIN AVE., SUITE 1100
CITY: MILWAUKEE
STATE: WISCONSIN
COUNTRY: USA
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,086
FILING DATE: 04-MAR-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: SARA, CHARLES S
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: C.8493-87
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 255-2022
TELEFAX: (608) 255-2182
TELEX: 26832 ANDSTARK
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 25
TYPE: Amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
FEATURE:
NAME/KEY: Phosphoserine
LOCATION: 15
OTHER INFORMATION: Post-translationally phosphorylated ser
OTHER INFORMATION: Post-translationally phosphorylated ser
FEATURE:
NAME/KEY: Phosphoserine
LOCATION: 18
OTHER INFORMATION: Post-translationally phosphorylated ser
OTHER INFORMATION: Post-translationally phosphorylated ser
FEATURE:
NAME/KEY: Phosphoserine
LOCATION: 19
OTHER INFORMATION: Post-translationally phosphorylated ser
OTHER INFORMATION: Post-translationally phosphorylated ser
US-08-137-086-1
Query Match 96.2%; Score 101; DB 4; Length 25;
Best Local Similarity 84.0%; Pred. No. 2,7e-10;
Matches 21; Conservative 0; Mismatches 4; Indels
QY 1 RELEELNVPGEIVEXLXXXEESITR 25
DB 1 RELEELNVPGEIVESLSSESISITR 25
RESULT 5
US-09-269-220-2
; Sequence 2, Application US/09269220
; Patent No. 6180761

```

```

, STREET: 711 Elizabeth Street
, CITY: Melbourne
, STATE: Victoria
, COUNTRY: Australia
, ZIP: 3000
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Diskette
, COMPUTER: TBM
, OPERATING SYSTEM: Windows 95
, SOFTWARE: Ascii
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/954,985A
, FILING DATE: 21-OCT-1997
, CLASSIFICATION:
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 08/362,479
, FILING DATE: 22-FEB-1995
, ATTORNEY/AGENT INFORMATION:
, NAME: Wozny, Thomas M
, REGISTRATION NUMBER: 28,922
, REFERENCE/DOCKET NUMBER: 322-00033
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (414) 271-7590
, TELEFAX: (414) 271-5770
, INFORMATION FOR SEQ ID NO: 3:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 24
, TYPE: Amino Acid
, TOPOLOGY: Linear
, FEATURE:
, NAME/KEY: Phosphoserine
, LOCATION: 14
, OTHER INFORMATION: Post-translation
, FEATURE:
, NAME/KEY: Phosphoserine
, LOCATION: 16
, OTHER INFORMATION: Post-translation
, FEATURE:
, NAME/KEY: Phosphoserine
, LOCATION: 17
, OTHER INFORMATION: Post-translation
, FEATURE:
, NAME/KEY: Phosphoserine
, LOCATION: 18
, OTHER INFORMATION: Post-translation
, OTHER INFORMATION: Post-translation
US-08-954-985A-3

Query Match 91.4%; Score 96;
Best Local Similarity 83.3%; Pred. No.
Matches 20; Conservative 0; Mismatch

QY 2 EELNVPGEIVEXLXXHSEISIR 25
Db 1 EELNVPGEIVESLSSSEISIR 24
|||||
|||||

RESULT 9
US-08-137-086-3
, Sequence 3, Application US/08137086
, Patent No. 6448374
, GENERAL INFORMATION:
, APPLICANT: REYNOLDS, ERIC CHARLES
, TITLE OF INVENTION: PRODUCTION OF PHOS
, NUMBER OF SEQUENCES: 9
, CORRESPONDENCE ADDRESS:
, ADDRESSEE: ANDRUS, SCEALES, STARKE &
, STREET: 100 EAST WISCONSIN AVE., SUITE
, CITY: MILWAUKEE
, STATE: WISCONSIN
, COUNTRY: USA
, ZIP: 53202
, COMPUTER READABLE FORM:
, MEDIUM TYPE: FLOPPY DISK

```

ZIP: 22314
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/621,564B
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Kraus, Eric J.
 REGISTRATION NUMBER: 36,190
 REFERENCE/DOCKET NUMBER: 2438-022
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 684-1111
 TELEFAX: (703) 684-1124
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ID: US-08-621-564B-1

Query Match	91.4%;	Score 96;	DB 2;	Length 28;
Best Local Similarity	83.3%;	Pred. No. 2.1e-09;		
Matches 20;	Conservative	0;	Mismatches 4;	Indels 0;
				Gaps 0;

1 RELEELNVPGEIVEXLXXEESIT 24
1 RELEELNVPGEIVEXLXXEESIT 24

```

RESULT 11
US-09-269-220-1
; Sequence 1, Application US/09269220
; Patent No. 6180761
; GENERAL INFORMATION:
; APPLICANT: HAN, Sang K
; APPLICANT: SHIN, Yoo C
; TITLE OF INVENTION: CASEIN AND PROCESS FOR THE PREPARATION THEREOF
; FILE REFERENCE: 1423.1001/MJH
; CURRENT APPLICATION NUMBER: US/09/269,220
; CURRENT FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: KR 1996-43482
; PRIOR FILING DATE: 1996-03-23
; PRIOR APPLICATION NUMBER: PCT/KR97/00182
; PRIOR FILING DATE: 1997-09-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: ACT_SITE
; LOCATION: (15)
; OTHER INFORMATION: phospholyated serine
; NAME/KEY: ACT_SITE
; LOCATION: (17)..(19)
; US-09-269-220-1

```

Query Match	91.4%;	Score 96;	DB 4;	Length 209;
Best Local Similarity	83.3%;	Pred. No. 2.1e-08;		
Conservative	0;	Mismatches 4;	Indels	

1 RELEELNPGEIVEIXLXXEESIT 24
1 RELEELNPGEIVEIXLXXEESIT 24

COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,086
FILING DATE: 04-MAR-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: SARA, CHARLES S
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: C.8493-87
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 255-2022
TELEFAX: (608) 255-2182
TELEX: 26832 ANDSTARK
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 24
TYPE: Amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
FEATURE:
NAME/KEY: Phosphoserine
LOCATION: 14
OTHER INFORMATION:
OTHER INFORMATION: Post-translationally phosphorylated serine
FEATURE:
NAME/KEY: Phosphoserine
LOCATION: 16
OTHER INFORMATION:
OTHER INFORMATION: Post-translationally phosphorylated serine
FEATURE:
NAME/KEY: Phosphoserine
LOCATION: 17
OTHER INFORMATION:
OTHER INFORMATION: Post-translationally phosphorylated serine
FEATURE:
NAME/KEY: Phosphoserine
LOCATION: 18
OTHER INFORMATION:
OTHER INFORMATION: Post-translationally phosphorylated serine
FEATURE:
NAME/KEY: Phosphoserine
LOCATION: 19
OTHER INFORMATION:
OTHER INFORMATION: Post-translationally phosphorylated serine

Query Match 91.4%; Score 96; DB 4; Length 24;
Best Local Similarity 83.3%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 EEEENVPGEIVEXLXXEESITR 25
 | | | | | | | | | | | | | |

DB 1 EEEENVPGEIVESLSSEESITR 24

```

RESULT 10
US-08-621-564B-1
; Sequence 1, Application US/08621564B
; Patent No. 5834427
; GENERAL INFORMATION:
; APPLICANT: Han, Sang Kee
; APPLICANT: Shin, Yoo Cheol
; TITLE OF INVENTION: CASEIN PHOSPHOPEPTIDE, CASEIN CONTAINING
; TITLE OF INVENTION: SAME, AND PROCESS FOR THE PREPARATION THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lowe, Price, LeBlanc & Becker
; STREET: 99 Canal Center Plaza, Suite 300
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA

```

```

RESULT 12
US-08-391-743A-2
; Sequence 2, Application US/08391743A
; Patent No. 5843705
; GENERAL INFORMATION:
; APPLICANT: DiTullio, Paul A.; Meade, Harry; Cole, Edward S.
; TITLE OF INVENTION: TRANSGENETICALLY PRODUCED ANTITHROMBIN III
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/391.743A
; FILING DATE: 21-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: TCI-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-391-743A-2

Query Match 72.4%; Score 76; DB 2; Length 222;
Best Local Similarity 70.8%; Pred. No. 4.3e-05;
Matches 17; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 RELEELNVPEIVEIXLXXEESIT 24
|| ||||| || || || |||||
Db 16 REQELNVVGETVESLSSESIT 39

RESULT 13
US-09-143-155-2
; Sequence 2, Application US/09143155
; Patent No. 6441145
; GENERAL INFORMATION:
; APPLICANT: DiTullio, Paul A.; Meade, Harry; Cole, Edward S.
; TITLE OF INVENTION: TRANSGENETICALLY PRODUCED ANTITHROMBIN III
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/143.155
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: TCI-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-143-155-2

Query Match 72.4%; Score 76; DB 2; Length 222;
Best Local Similarity 70.8%; Pred. No. 4.3e-05;
Matches 17; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 RELEELNVPEIVEIXLXXEESIT 24
|| ||||| || || || |||||
Db 16 REQELNVVGETVESLSSESIT 39

RESULT 14
US-09-131-028A-2
; Sequence 2, Application US/09131028A
; Patent No. 6287866
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Lemmel, Steven A.
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Chaudhary, Sunita
; TITLE OF INVENTION: BETA-CASEIN EXPRESSING CONSTRUCTS
; FILE REFERENCE: 6004.US.P1
; CURRENT APPLICATION NUMBER: US/09/131.028A
; CURRENT FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: US 08/064,440
; PRIOR FILING DATE: 1993-05-21
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-131-028A-2

Query Match 42.9%; Score 45; DB 4; Length 213;
Best Local Similarity 52.9%; Pred. No. 5.2;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 8 VPGEIVEXLXXEESIT 24
:|:|:|:|:|:|:|:|:|:|:|:|
Db 1 MPRETIESLSSESIT 17

RESULT 15
US-09-131-028A-12
; Sequence 12, Application US/09131028A
; Patent No. 6287866
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Lemmel, Steven A.
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Chaudhary, Sunita

```


us-09-380-738a-2.ra

Wed Feb-12 11:35:21 2003

;; TITLE OF INVENTION: BETA-CASEIN EXPRESSING CONSTRUCTS

;; FILE REFERENCE: 6004.US.P1

;; CURRENT APPLICATION NUMBER: US/09/131,028A

;; CURRENT FILING DATE: 1998-08-07

;; PRIOR APPLICATION NUMBER: US 08/064,440

;; PRIOR FILING DATE: 1993-05-21

;; NUMBER OF SEQ ID NOS: 22

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 12

;; LENGTH: 213

;; TYPE: PRT

;; ORGANISM: Homo sapiens

US-09-131-028A-12

Query Match 42.9%; Score 45; DB 4; Length 213;

Best Local Similarity 52.9%; Pred. No. 5.2;

Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

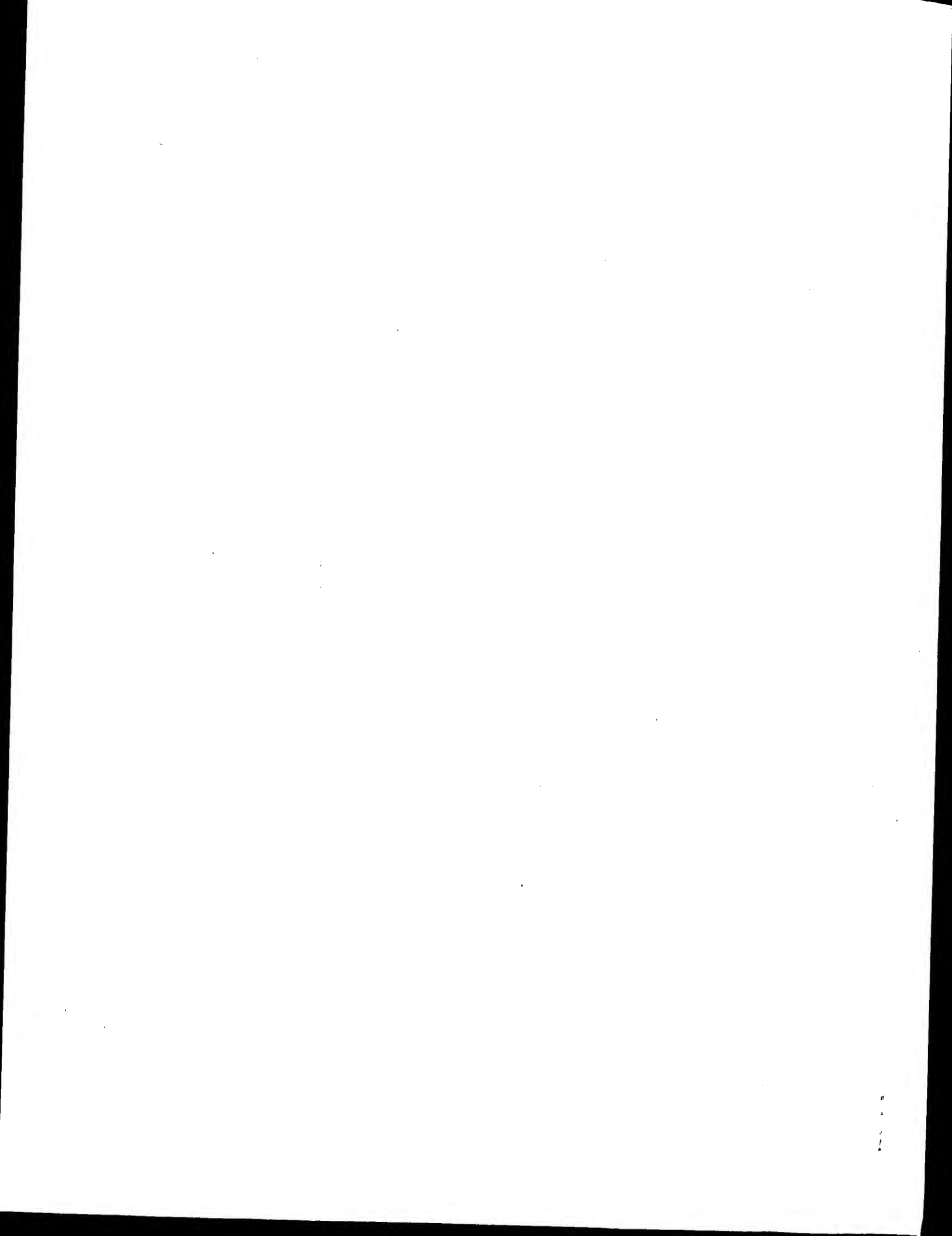
QY 8 VPGEIVEXLXXEESIT 24

:|:|:|:|:|:|

Db 1 MPRETIESLSSEESIT 17

Search completed: February 11, 2003, 18:22:08

Job time : 13.5711 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:19:51 ; Search time 12.1134 Seconds
(without alignments)
52.729 Million cell updates/sec

Title: US-09-380-738A-2

Perfect score: 105
Sequence: 1 RELELNVPGEIVLXXEESITR 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues
Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
 - 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
 - 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	43	41.0	74	10	US-09-110-716-38
2	43	41.0	112	10	US-09-934-054-8
3	43	41.0	2492	10	US-09-991-258-3
4	42	40.0	304	10	US-09-815-242-14025
5	41	39.0	275	9	US-09-738-626-5720
6	40	38.1	360	10	US-09-361-741-4
7	39	37.1	519	9	US-10-114-170-265
8	39	37.1	1248	9	US-09-738-626-4814
9	38	36.2	421	10	US-09-767-770A-4
10	38	36.2	611	9	US-09-893-519A-28
11	37	35.2	173	9	US-09-853-450-40
12	37	35.2	192	9	US-09-853-450-50
13	37	35.2	196	9	US-10-060-523-7
14	37	35.2	196	9	US-09-935-726-5
15	37	35.2	196	9	US-10-084-488-5
16	37	35.2	196	10	US-09-749-728B-3
17	37	35.2	196	10	US-09-795-006A-125
18	37	35.2	196	12	US-10-127-551-3
19	37	35.2	211	9	US-09-852-209A-14

20	37	35.2	389	9	US-09-712-363-196
21	37	35.2	786	9	US-10-045-792-11
22	37	35.2	793	9	US-10-045-792-10
23	37	35.2	1371	10	US-09-272-809-4
24	36	34.3	84	10	US-09-864-761-43315
25	36	34.3	224	10	US-09-904-568-2
26	36	34.3	387	9	US-10-114-893-133
27	36	34.3	521	9	US-09-884-566-2
28	36	34.3	621	12	US-10-042-417-28
29	36	34.3	640	10	US-09-918-951-4
30	36	34.3	642	10	US-09-867-550-1540
31	36	34.3	647	9	US-09-738-626-3644
32	36	34.3	750	10	US-09-815-242-13405
33	36	34.3	755	9	US-09-738-626-3935
34	36	34.3	760	10	US-09-866-582-42
35	36	34.3	765	9	US-10-217-357-4
36	36	34.3	765	10	US-09-975-326-4
37	36	34.3	766	9	US-09-934-406-2
38	36	34.3	766	9	US-10-217-357-2
39	36	34.3	766	10	US-09-975-326-2
40	36	34.3	794	10	US-09-815-242-5697
41	36	34.3	802	10	US-09-815-242-12668
42	36	34.3	1014	10	US-09-912-020-266
43	36	34.3	1226	10	US-09-815-242-13646
44	36	34.3	1242	10	US-09-925-299-911
45	36	34.3	1421	10	US-09-924-154-13

ALIGNMENTS

RESULT 1
US-09-110-716-38
; Sequence 38, Application US/09110716A
; Patent No. US20020034739A1
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Zhao, Chengquan
; APPLICANT: Glasgow, Benjamin J.
; TITLE OF INVENTION: PEPTIDES CHARACTERISTIC OF CERTAIN TUMORS
; FILE REFERENCE: 22000-20596.00
; CURRENT APPLICATION NUMBER: US/09/110,716A
; CURRENT FILING DATE: 1998-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 38
; TYPE: PRT
; ORGANISM: PRC2
US-09-110-716-38

Query Match 41.0%; Score 43; DB 10; Length 74;
Best Local Similarity 64.3%; Pred. No. 1.3;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy	1	RELELNVPGEIV 14
Db	23	RELEEFDPPEAVE 36

RESULT 2
US-09-934-054-8
; Sequence 8, Application US/09934054
; Patent No. US20020107385A1
; GENERAL INFORMATION:
; APPLICANT: Akerblom, Ingrid E.
; Hillman, Jennifer L.
; Murry, Lynn E.
; Goli, Surya K.
; Hawkins, Phillip R.
; TITLE OF INVENTION: BREAST TUMOR SPECIFIC PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025-6936
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/09/934,054
FILING DATE: 21-Aug-2001
PRIOR APPLICATION NUMBER: US/08/747,547
APPLICATION NUMBER: US/08/747,547
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0077 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 112 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-934-054-8

Query Match 41.0%; Score 43; DB 10; Length 112;
Best Local Similarity 64.3%; Pred. No. 2.2;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RELELNVPGEIVE 14
|||||: |||||
Db 48 RELEFDAPPEAVE 61

RESULT 3
US-09-991-258-3
; Sequence 3, Application US/09991258
; Patent No. US20020141975A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swanson, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
; FILE REFERENCE: 01113.0001U3
; CURRENT APPLICATION NUMBER: US/09/991,258
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2492
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020141975A1e =
; OTHER INFORMATION: Synthetic construct

US-09-991-258-3
Query Match 41.0%; Score 43; DB 10; Length 2492;
Best Local Similarity 40.9%; Pred. No. 82;
Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 4 EELNVPGEIVEXLXXEESITR 25
|||:||||: |||:
Db 1174 EKLSVPGKMDWLSRDPEATFR 1195

RESULT 4
US-09-815-242-14025
; Sequence 14025, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14025
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-14025

Query Match 40.0%; Score 42; DB 10; Length 304;
Best Local Similarity 41.2%; Pred. No. 10;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 9 PGEIVEXLXXEESITR 25
|||:||||: |||:
Db 21 PGKIDTLGAMQKSLTR 37

RESULT 5
US-09-738-626-5720
; Sequence 5720, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: YOTEISHI, NAOKO
; APPLICANT: SENO, AKIHIRO

us-09-380-738a-2.rapb

Wed Feb 12 11:35:21 2003

US-09-361-741-4

APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5720
LENGTH: 275
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5720

Query Match 38.1%; Score 40; DB 10; Length 360;
Best Local Similarity 54.5%; Pred. No. 27;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 LEELNVPEIV 13
:|| |||| :
Db 73 VEEFNPVGSVI 83

RESULT 7

US-10-114-170-265
; Sequence 265, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. US20030023075A1el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/114,170
FILING DATE: 01-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/453,702
FILING DATE: 03-DEC-1999
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166

INFORMATION FOR SEQ ID NO: 265:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 265:
US-10-114-170-265

Query Match 37.1%; Score 39; DB 9; Length 519;
Best Local Similarity 26.1%; Pred. No. 62;
Matches 6; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 REELNVPEIVEXLXXEESI 23
| | | | | :
Db 91 RSLVEVSIPEGLTDFRQIONNV 113

RESULT 8

US-09-738-626-4814
; Sequence 4814, Application US/09738626

APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5720
LENGTH: 275
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5720

Query Match 39.0%; Score 41; DB 9; Length 275;
Best Local Similarity 47.4%; Pred. No. 14;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 7 NVPEIVEXLXXEESITR 25
:|| |||| :
Db 196 DVPAEIEEKERSIAEQITR 214

RESULT 6

US-09-361-741-4
; Sequence 4, Application US/09361741
; Patent No. US20020048784A1
; GENERAL INFORMATION:
; APPLICANT: HUDSON, PETER L
; APPLICANT: ROSEN, CRAIG A
; APPLICANT: HE, WEI WU
; TITLE OF INVENTION: PROSTATIC GROWTH FACTOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NJ
; COUNTRY: US
; ZIP: 07068

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/361,741
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,607
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-329
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201)994-1700
TELEFAX: (201) 994-1744
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 360 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match 36.1%; Score 36; DB 134
Best Local Similarity 36.1%; Pred. No. 134
Matches 325800-329

QY 4
:|| |||| :
Db 4

Publication No. US20020197605A1
 GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIKO
 APPLICANT: HAYASHI, MIKIO
 APPLICANT: OCHIAI, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TATEISHI, NAOKO
 APPLICANT: SENOH, AKIHIRO
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738,626
 CURRENT FILING DATE: 2000-12-18
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO 4814
 LENGTH: 1248
 TYPE: PRT
 ORGANISM: Corynebacterium glutamicum
 US-09-738-626-4814

Query Match 37.1%; Score 39; DB 9; Length 1248;
 Best Local Similarity 87.5%; Pred. No. 1.7e+02;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 EEINVPGE 11
 Db 496 EEINLPGE 503
 |||:||||

RESULT 9
 US-09-767-770A-4
 Sequence 4, Application US/09767770A
 Patent No. US20020065237A1
 GENERAL INFORMATION:
 APPLICANT: Michalovich, David
 APPLICANT: Shaikh, Narjis
 APPLICANT: Sims, Matthew Alan
 TITLE OF INVENTION: NOVEL COMPOUNDS
 FILE REFERENCE: GP-30086-C1
 CURRENT APPLICATION NUMBER: US/09/767,770A
 CURRENT FILING DATE: 2001-01-23
 PRIOR APPLICATION NUMBER: 9806222.7
 PRIOR FILING DATE: 1998-03-23
 PRIOR APPLICATION NUMBER: 9820299.7
 PRIOR FILING DATE: 1998-09-17
 PRIOR APPLICATION NUMBER: 09/266,261
 PRIOR FILING DATE: 1999-03-11
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 4
 LENGTH: 421
 TYPE: PRT
 ORGANISM: HOMO SAPIENS
 US-09-767-770A-4

Query Match 36.2%; Score 38; DB 10; Length 421;
 Best Local Similarity 37.5%; Pred. No. 71;
 Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 RELEENVPGEIVEXL 16
 Db 402 QEVQELPIPSKLLLEFL 417
 |:|:|:|:|:|:|:|

RESULT 10
 US-09-893-519A-28
 Sequence 28, Application US/09893519A
 Publication No. US20030027243A1
 GENERAL INFORMATION:
 APPLICANT: ANADYS PHARMACEUTICALS, INC.
 APPLICANT: THOMPSON, Craig
 APPLICANT: MOORE, Jeffrey
 APPLICANT: BUURMAN, Ed T.
 APPLICANT: BRADLEY, John
 APPLICANT: DESILVA, Thamara
 APPLICANT: HARRIS, Sandra
 APPLICANT: KOMARNITSKY, Svetlana
 APPLICANT: MENDILLO, Marc
 APPLICANT: MOORE, Daniel
 APPLICANT: MCCOY, Melissa
 APPLICANT: SANDERSON, Karen
 APPLICANT: HAQ, Tariq
 APPLICANT: ZHU, Shuhao
 APPLICANT: LONG, Fan
 APPLICANT: DAVIDOV, Eugene
 TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
 FILE REFERENCE: 0342/1G548-US2
 CURRENT APPLICATION NUMBER: US/09/893,519A
 CURRENT FILING DATE: 2001-06-28
 PRIOR APPLICATION NUMBER: US 60/215,164
 PRIOR FILING DATE: 2000-06-29
 PRIOR APPLICATION NUMBER: US 60/224,457
 PRIOR FILING DATE: 2000-08-10
 NUMBER OF SEQ ID NOS: 146
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 28
 LENGTH: 611
 TYPE: PRT
 ORGANISM: Homo sapiens
 NAME/KEY: misc_feature
 OTHER INFORMATION: Corresponds to SEQ ID NO: 101
 PUBLICATION INFORMATION:
 DATABASE ACCESSION NUMBER: Human Genbank/NP_001078
 DATABASE ENTRY DATE: 2001-12-18
 RELEVANT RESIDUES: (1)..(611)
 US-09-893-519A-28

Query Match 36.2%; Score 38; DB 9; Length 611;
 Best Local Similarity 37.5%; Pred. No. 1.1e+02;
 Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 RELEENVPGEIVEXL 16
 Db 592 QEVQELPIPSKLLLEFL 607
 |:|:|:|:|:|:|

RESULT 11
 US-09-853-450-40
 Sequence 40, Application US/09853450
 Publication No. US20020194645A1
 GENERAL INFORMATION:
 APPLICANT: Yanofsky, Martin F.
 APPLICANT: Pelaz, Soraya
 APPLICANT: Ditta, Gary
 APPLICANT: The Regents of the University of California
 TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
 FILE OF INVENTION: Exhibiting Modulated Reproductive Development
 FILE REFERENCE: 19452A-002400US
 CURRENT APPLICATION NUMBER: US/09/853,450
 CURRENT FILING DATE: 2001-05-09
 NUMBER OF SEQ ID NOS: 61
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 40
 LENGTH: 173
 TYPE: PRT

```

RESULT 13
US-10-060-523-7
; Sequence 7, Application US/10060523
; Publication No. US2002018263A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Jing-Shan
; APPLICANT: Rosen, Craig A.
; APPLICANT: Cao, Liang
; TITLE OF INVENTION: Polynucleotide
; FILE REFERENCE: Pfl12D1
; CURRENT APPLICATION NUMBER: US/10-060-523-7
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US/09/61
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 09/257,9
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 08/824,9
; PRIOR FILING DATE: 1997-03-27
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-523-7

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

NAME: MICHELE M. WALES
REGISTRATION NUMBER: 43,975
REFERENCE/DOCKET NUMBER: PFI12PCT3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301)309-8504
TELEFAX: (301)309-8439
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 196 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-084-488-5

```

Query Match          35.2%;      Score 37;  DB 9;  Length 196;
Best Local Similarity 35.3%;
Matches             6;  Conservative 4;  Mismatches 7;  Indels 0; Gaps 0;

```

Search completed: February 11, 2003, 18:36:14
Job time : 13.1134 secs

GenCore version 5.1.3
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: February 11, 2003, 18:17:01 : Search time 171.649 Seconds
(without alignments)
93.903 Million cell updates/sec

Title: US-09-380-738A-2
Perfect score: 105
Sequence: 1 RELELNVPGEIVEXLXXEESITR 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 4569144 seqs, 644733110 residues
Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	101	96.2	25	10 US-08-621-564-2	Sequence 2, Appli
2	101	96.2	25	10 US-08-621-564A-2	Sequence 2, Appli
3	101	96.2	25	16 US-09-222-791-2	Sequence 23, Appl
4	101	96.2	48	10 US-08-666-559-33	Sequence 33, Appl
5	101	96.2	209	10 US-08-666-559-23	Sequence 2, Appli
6	97	92.4	25	17 US-09-380-738A-2	Sequence 2, Appli

7	96	91.4	24	3	US-07-731-592B-2	Sequence 2, Appli
8	96	91.4	28	10	US-08-621-564-1	Sequence 1, Appli
9	96	91.4	28	10	US-08-621-564A-1	Sequence 1, Appli
10	85.5	81.4	23	10	US-08-666-559-103	Sequence 103, App
11	55	52.4	12	10	US-08-666-559-35	Sequence 35, Appl
12	51	48.6	321	26	US-10-219-999-52212	Sequence 52212, A
13	46	43.8	334	21	US-09-708-427-34394	Sequence 34394, A
14	45	42.9	488	21	US-09-791-537-141484	Sequence 141484, A
15	45	42.9	492	21	US-09-791-537-141465	Sequence 141465, A
16	44.5	42.4	687	26	US-10-219-999-32152	Sequence 32152, A
17	44.5	42.4	687	27	US-60-324-109-17136	Sequence 17126, A
18	44.5	42.4	689	25	US-10-155-881-11750	Sequence 11750, A
19	43	41.0	74	15	US-09-110-716-38	Sequence 38, Appli
20	43	41.0	112	11	US-08-747-547-8	Sequence 8, Appli
21	43	41.0	112	23	US-09-934-054-8	Sequence 92101, A
22	43	41.0	472	21	US-09-791-537-92101	Sequence 3353, A
23	43	41.0	526	26	US-10-219-999-33353	Sequence 18580, A
24	43	41.0	526	27	US-60-324-109-18580	Sequence 8064, Ap
25	43	41.0	660	16	US-09-252-691-8064	Sequence 8064, Ap
26	43	41.0	660	16	US-09-252-691C-8064	Sequence 2083, Ap
27	43	41.0	678	19	US-09-573-655B-2083	Sequence 2083, Ap
28	43	41.0	974	21	US-09-791-537-92019	Sequence 92019, A
29	43	41.0	2492	1	PCT-US01-21701-3	Sequence 3, Appli
30	43	41.0	2492	23	US-09-902-537-3	Sequence 3, Appli
31	43	41.0	2492	23	US-09-991-258-3	Sequence 886, App
32	43	41.0	140	1	PCT-US98-06371-886	Sequence 152, App
33	42.5	40.5	140	13	US-08-902-615A-152	Sequence 61885, A
34	42.5	40.5	140	21	US-09-791-537-61885	Sequence 152, App
35	42.5	40.5	140	22	US-09-882-227-152	Sequence 15290, A
36	42.5	40.5	198	16	US-09-248-796-15290	Sequence 15290, A
37	42.5	40.5	198	27	US-60-096-409-15290	Sequence 8159, Ap
38	42.5	40.5	522	25	US-10-179-131-8159	Sequence 2147, Ap
39	42.5	40.5	522	25	US-60-161-932-2147	Sequence 32346, A
40	42.5	40.5	262	1	PCT-US01-08631-32346	Sequence 14025, A
41	42	40.0	304	1	PCT-US02-03987-14025	Sequence 14025, A
42	42	40.0	304	22	US-09-815-242-14025	Sequence 14025, A
43	42	40.0	304	24	US-10-072-851-14025	Sequence 10621, A
44	42	40.0	304	24	US-10-072-851-14025	Sequence 10621, A
45	42	40.0	367	23	US-09-902-540-10621	Sequence 10621, A

ALIGNMENTS

RESULT 1
US-08-621-564-2
; Sequence 2, Application US/08621564
; GENERAL INFORMATION:
; APPLICANT: Han, Sang Kee
; APPLICANT: Shin, Yoo Cheol
; TITLE OF INVENTION: CASEIN PHOSPHOPEPTIDE, CASEIN CONTAINING
; TITLE OF INVENTION: SAME, AND PROCESS FOR THE PREPARATION THEREOF
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lowe, Price, Leblanc & Becker
; STREET: 99 Canal Center Plaza, Suite 300
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/621,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Kraus, Eric J.
; REGISTRATION NUMBER: 36,190
; REFERENCE/DOCKET NUMBER: 2438-022

```
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 684-1111
; TELEFAX: (703) 684-1124
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-621-564-2

Query Match          96.2%; Score 101; DB 10; Length 25;
Best Local Similarity 84.0%; Pred. No. 5.5e-10;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RELEELNPGEIVEXLXXEESITR 25
Db 1 RELEELNPGEIVESLSSEESITR 25

RESULT 2
US-08-621-564A-2
; Sequence 2, Application US/08621564A
; GENERAL INFORMATION:
; APPLICANT: HAN, Sang Kee
; TITLE OF INVENTION: CASEIN PHOSPHOPEPTIDE, CASEIN CONTAINING
; TITLE OF INVENTION: SAME AND PROCESS FOR THE PREPARATION THEREOF
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LOWE PRICE LEBLANC & BECKER
; STREET: 99 Canal Center Plaza, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/621,564A
; FILING DATE: 25-MAR-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Hauptman, Benjamin J.
; REGISTRATION NUMBER: 29,310
; REFERENCE/DOCKET NUMBER: 2438-022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-684-1111
; TELEFAX: 703-684-1124
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-621-564A-2

Query Match          96.2%; Score 101; DB 10; Length 25;
Best Local Similarity 84.0%; Pred. No. 5.5e-10;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RELEELNPGEIVEXLXXEESITR 25
Db 1 RELEELNPGEIVESLSSEESITR 25

RESULT 3
US-09-222-791-2

; Sequence 2, Application US/09222791
; GENERAL INFORMATION:
; APPLICANT: Chen, Wen-Yih
; APPLICANT: Huang, Shih-Yow
; APPLICANT: Lin, Fu-Yong
; TITLE OF INVENTION: A Method for Separating and Purifying
; TITLE OF INVENTION: Caseinophosphopeptides from Casein Hydrolysate by Using
; FILE REFERENCE: Serial No. 09/222,791
; CURRENT APPLICATION NUMBER: US/09/222,791
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-222-791-2

Query Match          96.2%; Score 101; DB 16; Length 25;
Best Local Similarity 84.0%; Pred. No. 5.5e-10;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RELEELNPGEIVEXLXXEESITR 25
Db 1 RELEELNPGEIVESLSSEESITR 25

RESULT 4
US-08-666-559-33
; Sequence 33, Application US/08666559
; GENERAL INFORMATION:
; APPLICANT: Jean-Marc CHOBERT et al.
; TITLE OF INVENTION: METHOD FOR PREPARING PEPTIDE
; TITLE OF INVENTION: PRODUCTS, AND RESULTING PRODUCTS
; NUMBER OF SEQUENCES: 122
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenblum & Bernstein, P.L.C.
; STREET: 1941 Roland Clarke Place
; CITY: Reston
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/666,559
; FILING DATE: September 10, 1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 93 15764
; FILING DATE: December 23, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold Turk
; REGISTRATION NUMBER: 33,094
; REFERENCE/DOCKET NUMBER: P14764
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 716-1191
; TELEFAX: (703) 716-1180
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal fragment
US-08-666-559-33

Query Match          96.2%; Score 101; DB 10; Length 48;
```

```

: Sequence 2, Application US/09380/39A
:
: GENERAL INFORMATION:
:
: APPLICANT: REYNOLDS, Eric
: TITLE OF INVENTION: CALCIUM PHOSPHOPEPTIDE COMPLEXES
: FILE REFERENCE: 040268/0161
: CURRENT APPLICATION NUMBER: US/09/380,738A
: CURRENT FILING DATE: 1999-12-06
: PRIOR APPLICATION NUMBER: PCT/AU98/00160
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: AU P05662

```

```
, , LOCATION: 14 post-translation  
, , OTHER INFORMATION: serine  
, , OTHER INFORMATION:  
, , FEATURE:  
, , NAME/KEY: phosphoserine  
, , LOCATION: 16 Post-translation  
, , OTHER INFORMATION: serine  
, , OTHER INFORMATION:
```

;
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 17
; OTHER INFORMATION: Post-translationally phosphorylated
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 18
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
US-07-731-592B-2

Query Match 91.4%; Score 96; DB 3; Length 24;
Best Local Similarity 83.3%; Pred. No. 3.9e-09;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 EELELNPGEIVEXLXXEESITR 25
Db 1 EELELNPGEIVESLSSESITR 24

RESULT 8

US-08-621-564-1
; Sequence 1, Application US/08621564
; GENERAL INFORMATION:
; APPLICANT: Han, Sang Kee
; TITLE OF INVENTION: CASEIN PHOSPHOPEPTIDE, CASEIN CONTAINING
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lowe, Price, Leblanc & Becker
; STREET: 99 Canal Center Plaza, Suite 300
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 25-MAR-1996
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Kraus, Eric J.
REGISTRATION NUMBER: 36,190
REFERENCE/DOCKET NUMBER: 2438-022
TELEPHONE: (703) 684-1111
TELEFAX: (703) 684-1124
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-621-564-1

Query Match 91.4%; Score 96; DB 10; Length 28;
Best Local Similarity 83.3%; Pred. No. 4.8e-09;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RELELNPGEIVEXLXXEESIT 24
Db 1 RELELNPGEIVESLSSESIT 24

RESULT 9

US-08-621-564A-1

; Sequence 1, Application US/08621564A
; GENERAL INFORMATION:
; APPLICANT: Han, Sang Kee
; TITLE OF INVENTION: CASEIN PHOSPHOPEPTIDE, CASEIN CONTAINING
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LOWE PRICE LEBLANC & BECKER
; STREET: 99 Canal Center Plaza, Suite 300
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 25-MAR-1996
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Hauptman, Benjamin J.
REGISTRATION NUMBER: 29,310
REFERENCE/DOCKET NUMBER: 2438-022
TELEPHONE: 703-684-1111
TELEFAX: 703-684-1124

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-621-564A-1

Query Match

Best Local Similarity 91.4%; Score 96; DB 10; Length 28;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RELELNPGEIVEXLXXEESIT 24
Db 1 RELELNPGEIVESLSSESIT 24

RESULT 10

US-08-666-559-103
; Sequence 103, Application US/08666559
; GENERAL INFORMATION:
; APPLICANT: Jean-Marc CHOBERT et al.
; TITLE OF INVENTION: METHOD FOR PREPARING PEPTIDE
; NUMBER OF SEQUENCES: 122
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenblum & Bernstein, P.L.C.
; STREET: 1941 Roland Clarke Place
; CITY: Reston
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 20191

COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Perfect 5.1+ (ASCII)
CURRENT APPLICATION DATA:
FILING DATE: September 10, 1996
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
US-08-666-559-103

Query Match 52.4%; Score 55; DB 10; Length 12;

QY 1 RELEELNVPGEIVEXLXXX--EESITR 25
:||||:| ||| :
:|:|

Db 32 KEFEKIHIPNEIVEILVRLPVKSLTR 58

RESULT 14

US-09-791-537-141484
; Sequence 141484, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 141484
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Methanobacterium thermoautotrophicum
US-09-791-537-141484

Query Match 42.9%; Score 45; DB 21; Length 488;
Best Local Similarity 61.5%; Pred. No. 1.9e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 EELNVPGEIVEXL 16
:|||||: :|
Db 201 KELNVPGDAIEKL 213

RESULT 15

US-09-791-537-141465
; Sequence 141465, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 141465
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Methanobacterium thermoautotrophicum
US-09-791-537-141465

Query Match 42.9%; Score 45; DB 21; Length 492;
Best Local Similarity 61.5%; Pred. No. 2e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 EELNVPGEIVEXL 16
:|||||: :|
Db 205 KELNVPGDAIEKL 217

Search completed: February 11, 2003, 18:33:21
Job time : 172.649 secs

Wed Feb 12 11:35:22 2003

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: February 11, 2003, 18:17:56 ; search time 28.6082 Seconds
(without alignments)
72.734 Million cell updates/sec

Title: US-09-380-738A-2

Perfect score: 105
Sequence: 1 RELELNVPGEIVEXLXXEESITR 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 420401 seqs, 83231269 residues
Total number of hits satisfying chosen parameters: 420401

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA.New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	41	39.0	196	US-09-857-346-27
2	41	39.0	221	PCT-US02-32727-22619
3	41	39.0	221	Sequence 22619, A
4	41	39.0	1037	Sequence 22619, A
5	40	38.1	360	Sequence 4794, A
6	39	37.1	1067	Sequence 4, Appli
7	39	37.1	1113	Sequence 5460, Ap
8	38	36.2	82	Sequence 8, Appli
9	38	36.2	82	Sequence 68459, A
10	38	36.2	196	Sequence 28, Appl
11	38	36.2	270	Sequence 5016, Ap
12	38	36.2	701	Sequence 15497, A
13	38	36.2	701	Sequence 15497, A
14	38	36.2	754	Sequence 4, Appli
15	38	36.2	833	Sequence 3, Appli
16	38	36.2	1752	Sequence 294, App
17	38	36.2	1752	Sequence 17, Appl
18	38	36.2	1805	Sequence 9319, Ap
19	38	36.2	1822	Sequence 485, Ap
20	38	36.2	1917	Sequence 4, Appli
21	37.5	35.7	188	Sequence 632, App
22	37.5	35.7	1799	Sequence 7957, Ap
23	37	35.2	83	Sequence 7957, Ap
24	37	35.2	100	Sequence 5858, Ap
25	37	35.2	113	Sequence 7281, Ap
26	37	35.2	113	Sequence 7281, Ap

Sequence 29941, A
Sequence 3821, Ap
Sequence 3821, Ap
Sequence 40, Appl
Sequence 61154, A
Sequence 61154, A
Sequence 50, Appl
Sequence 5, Appl
Sequence 26, Appl
Sequence 1, Appl
Sequence 5503, Ap
Sequence 5503, Ap
Sequence 61152, A
Sequence 61152, A
Sequence 61153, A
Sequence 61153, A
Sequence 5439, Ap
Sequence 23, Appl
Sequence 1841, Ap

ALIGNMENTS

RESULT 1
US-09-857-346-27 Application US/09857346

GENERAL INFORMATION:
APPLICANT: Burn, Joanne E
APPLICANT: Peacock, William J
APPLICANT: Dennis, Elizabeth S
APPLICANT: Sheldon, Candice C
APPLICANT: Helliwell, Christopher A
APPLICANT: Rouse, Dean T
TITLE OF INVENTION: Control of Flowering
FILE REFERENCE: 050341-0041
CURRENT APPLICATION NUMBER: US/09/857,346
CURRENT FILING DATE: 2002-06-04
PRIOR APPLICATION NUMBER: PP 7469
PRIOR FILING DATE: 1998-12-03
PRIOR APPLICATION NUMBER: US 60/116,928
PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: PCT/AU99/01079
PRIOR FILING DATE: 1999-12-02
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn version 3.1
SEQ ID NO 27
LENGTH: 196
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-857-346-27

Query Match 39.0%; Score 41; DB 5; Length 196;
Best Local Similarity 40.0%; Pred. No. 26;
Matches 12; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

Oy 2 ELELNVPGEIVEXLXXE-----SITR 25
Db 103 KLEESVNDASVDTLISLEQLETALSVTR 132

RESULT 2
PCT-US02-32727-22619

Sequence 22619, Application PC/TUS0232727
GENERAL INFORMATION:
APPLICANT: Mitcham, Jennifer
APPLICANT: Skeiky, Yasir
APPLICANT: Persing, David
APPLICANT: Bhatia, Ajay
APPLICANT: Maisonneuve, Jean Francois
APPLICANT: Zhang, Yanni
APPLICANT: Wang, Shiqian
APPLICANT: Jen, Shyan

SULT 4
-10-092-411A-4794
Sequence 4794, Application US/10092411A
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/10/092,411A
PRIORITY FILING DATE: 2002-03-07
PRIORITY APPLICATION NUMBER: US 09/134,001
PRIORITY FILING DATE: 1998-08-13
PRIORITY APPLICATION NUMBER: US 60/064,964
PRIORITY FILING DATE: 1997-11-08
PRIORITY APPLICATION NUMBER: US 60/055,779
PRIORITY FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5676
EQ ID NO 4794
LENGTH: 1037
TYPE: PRT

LENGTH: 103/
TYPE: PRT

PRIOR APPLICATION NUMBER: US 09/

PRIOR FILING DATE: 1999

PRIOR APPLICATION NUMBER

us-09-380-738a-2.rapn

wed Feb 12 11:35:22 2003

; PRIOR FILING DATE: 1998-03-10
 ; PRIOR APPLICATION NUMBER: US 09/036,720
 ; PRIOR FILING DATE: 1998-03-06
 ; PRIOR APPLICATION NUMBER: US 09/036,338
 ; PRIOR FILING DATE: 1998-03-06
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 7451
 ; SEQ ID NO 5460
 ; LENGTH: 1067
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 ; US-09-950-084-5460

Query Match 37.1%; Score 39; DB 5; Length 1067;
 Best Local Similarity 29.2%; Pred. No. 4.8e+02;
 Matches 7; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 2 ELEELNVPGEIVEXLXXEESITR 25
 Db 983 KLAEDIPAEVVGKIGENDLLTR 1006

RESULT 7
 US-09-332-522E-8
 ; Sequence 8, Application US/09332522E
 ; GENERAL INFORMATION:
 ; APPLICANT: Costa, M.
 ; APPLICANT: Doberstein, S.
 ; APPLICANT: Elson, S.
 ; APPLICANT: Ferguson, K.
 ; APPLICANT: Homberger, S.
 ; TITLE OF INVENTION: ANIMAL MODELS AND METHODS FOR ANALYSIS OF LIPID METABOLISM AND S
 ; TITLE OF INVENTION: OF PHARMACEUTICAL AND PESTICIDAL AGENTS THAT MODULATE LIPID METAB
 ; FILE REFERENCE: 7326-101, EX99-004
 ; CURRENT APPLICATION NUMBER: US/09/332,522E
 ; CURRENT FILING DATE: 1999-06-14
 ; NUMBER OF SEQ ID NOS: 95
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 1113
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 ; US-09-332-522E-8

Query Match 37.1%; Score 39; DB 5; Length 1113;
 Best Local Similarity 42.9%; Pred. No. 5e+02;
 Matches 9; Conservative 3; Mismatches 1; Indels 8; Gaps 1;

QY 2 ELEELN-----VPGEIVE 14
 Db 955 ELEHMMNEDDGFYQVPEVLE 975

RESULT 8
 US-09-724-676-68459
 ; Sequence 68459, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 68459
 ; LENGTH: 82
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-724-676-68459

Query Match 36.2%; Score 38; DB 5; Length 82;
 Best Local Similarity 50.0%; Pred. No. 28;
 Matches 7; Conservative 4; Mismatches 3; Indels 3;

QY 1 RELEELNVPGEIVE 14
 Db 34 RGLQETDVPFALVD 47

RESULT 9
 US-09-724-676A-68459
 ; Sequence 68459, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 68459
 ; LENGTH: 82
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-724-676A-68459

Query Match 36.2%; Score 38; DB 5; Length 82;
 Best Local Similarity 50.0%; Pred. No. 28;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 RELEELNVPGEIVE 14
 Db 34 RGLQETDVPFALVD 47

RESULT 10
 US-09-857-346-28
 ; Sequence 28, Application US/09857346
 ; GENERAL INFORMATION:
 ; APPLICANT: Burn Joanne E
 ; APPLICANT: Peacock, William J.
 ; APPLICANT: Dennis, Elizabeth S
 ; APPLICANT: Sheldon, Candice C.
 ; APPLICANT: Helliwell, Christopher A.
 ; APPLICANT: Rouse, Dean T.
 ; TITLE OF INVENTION: Control of Flowering
 ; FILE REFERENCE: 050341-0041
 ; CURRENT APPLICATION NUMBER: US/09/857,346
 ; CURRENT FILING DATE: 2002-06-04
 ; PRIOR APPLICATION NUMBER: PP 7469
 ; PRIOR FILING DATE: 1998-12-03
 ; PRIOR APPLICATION NUMBER: US 60/116,928
 ; PRIOR FILING DATE: 1999-01-22
 ; PRIOR APPLICATION NUMBER: PCT/AU99/01079
 ; PRIOR FILING DATE: 1999-12-02
 ; NUMBER OF SEQ ID NOS: 42
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 28
 ; LENGTH: 196
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-857-346-28

Query Match 36.2%; Score 38; DB 5; Length 196;
 Best Local Similarity 40.9%; Pred. No. 84;
 Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 2 ELEELNVPGEIVEXLXXEESI 23
 Db 103 KLEESVDNVSDSLISMEEQL 124

RESULT 11
 US-10-092-411A-5016
 ; Sequence 5016, Application US/10092411A
 ; GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 FILE REFERENCE: 210121.514
 CURRENT APPLICATION NUMBER: US/10/092,411a
 CURRENT FILING DATE: 2002-03-07
 PRIOR FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/134,001
 PRIOR FILING DATE: 1997-11-08
 PRIOR APPLICATION NUMBER: US 60/064,964
 PRIOR FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 5676
 SEQ ID NO 5016
 LENGTH: 270
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis
 US-10-092-411a-5016

Query Match 36.2%; Score 38; DB 6; Length 270;
 Best Local Similarity 53.8%; Pred. No. 1.3e+02;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ELEELNVPGEIVE 14
 Db 31 QLESFKPGEILE 43

RESULT 12
 PCT-US02-32727-15497
 Sequence 15497, Application PC/TUS0232727
 GENERAL INFORMATION:
 APPLICANT: Mitcham, Jennifer
 APPLICANT: Skeiky, Yasir
 APPLICANT: Persing, David
 APPLICANT: Bhatia, Ajay
 APPLICANT: Maisonneuve, Jean Francois
 APPLICANT: Zhang, Yanni
 APPLICANT: Wang, Siqing
 APPLICANT: Jen, Shyian
 APPLICANT: Lodes, Michael
 APPLICANT: Benson, Darin
 APPLICANT: Jones, Robert
 APPLICANT: Carter, Darriek
 APPLICANT: Barth, Brenda
 APPLICANT: Douglass, John

TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
 FILE REFERENCE: 210121.514C1
 CURRENT APPLICATION NUMBER: PCT/US02/32727
 CURRENT FILING DATE: 2002-10-11
 NUMBER OF SEQ ID NOS: 30992
 SEQ ID NO 15497
 LENGTH: 701
 TYPE: PRT
 ORGANISM: Propioni acnes
 PCT-US02-32727-15497

Query Match 36.2%; Score 38; DB 1; Length 701;
 Best Local Similarity 31.6%; Pred. No. 4.2e+02;
 Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 7 NVPGEIVEXLXXEESITR 25
 Db 483 DISGHLIQRVHRHEAITR 501

RESULT 13
 US-10-057-498-15497
 Sequence 15497, Application US/10057498
 GENERAL INFORMATION:
 APPLICANT: Mitcham, Jennifer
 APPLICANT: Skeiky, Yasir
 APPLICANT: Persing, David

TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Ac
 FILE REFERENCE: 210121.514
 CURRENT APPLICATION NUMBER: US/10/057,498
 CURRENT FILING DATE: 2001-04-20
 NUMBER OF SEQ ID NOS: 29212
 SEQ ID NO 15497
 LENGTH: 701
 TYPE: PRT
 ORGANISM: Propioni acnes
 US-10-057-498-15497

Query Match 36.2%; Score 38; DB 6; Length 701;
 Best Local Similarity 31.6%; Pred. No. 4.2e+02;
 Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 7 NVPGEIVEXLXXEESITR 25
 Db 483 DISGHLIQRVHRHEAITR 501

RESULT 14
 US-10-259-453-4
 Sequence 4, Application US/10259453
 GENERAL INFORMATION:
 APPLICANT: COMPUGEN LTD
 APPLICANT: AZAR, Idit
 APPLICANT: LEVINE, Zurit
 APPLICANT: BERNSTEIN, Jeanne
 APPLICANT: SAVITZKY, Kinnetet
 TITLE OF INVENTION: P13K-REGULATORY SUBUNIT HOMOLOGY
 FILE REFERENCE: 2786-0226P
 CURRENT APPLICATION NUMBER: US/10/259,453
 CURRENT FILING DATE: 2003-01-14
 PRIOR APPLICATION NUMBER: IL 135310
 PRIOR FILING DATE: 2000-03-28
 NUMBER OF SEQ ID NOS: 5
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 4
 LENGTH: 754
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-259-453-4

Query Match 36.2%; Score 38; DB 6; Length 754;
 Best Local Similarity 30.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 5 ELNVPGEIVEXLXXEESIT 24
 Db 126 EMAVEGTLYQRMVIAEQNL 145

RESULT 15
 US-10-259-453-3
 Sequence 3, Application US/10259453
 GENERAL INFORMATION:
 APPLICANT: COMPUGEN LTD
 APPLICANT: AZAR, Idit
 APPLICANT: LEVINE, Zurit
 APPLICANT: BERNSTEIN, Jeanne
 APPLICANT: SAVITZKY, Kinnetet
 TITLE OF INVENTION: P13K-REGULATORY SUBUNIT HOMOLOGY
 FILE REFERENCE: 2786-0226P
 CURRENT APPLICATION NUMBER: US/10/259,453
 CURRENT FILING DATE: 2003-01-14
 PRIOR APPLICATION NUMBER: IL 135310
 PRIOR FILING DATE: 2000-03-28
 NUMBER OF SEQ ID NOS: 5
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 3
 LENGTH: 833
 TYPE: PRT
 ORGANISM: Homo sapiens

us-09-380-738a-2.rapn

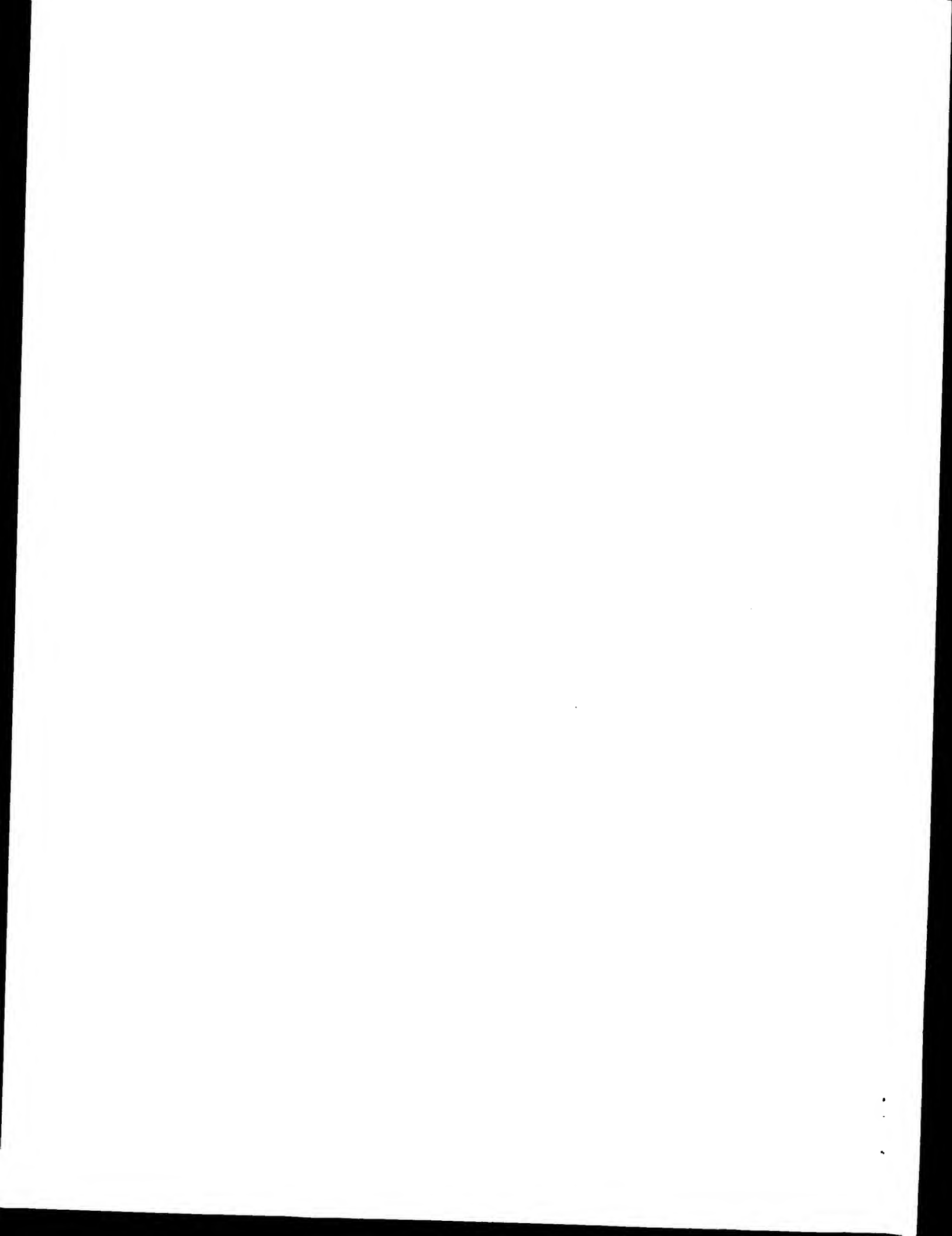
wed Feb' 12 11:35:22 2003

US-10-259-453-3

Query Match 36.2%; Score 38; DB 6; Length 833;
 Best Local Similarity 30.0%; Pred. No. 5.2e+02;
 Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 5 ELNVPGEIVEIXLXXEESIT 24
 I: ||| : : |::|
 Db 205 EMVPTLYQRMVIAEQNLT 224

Search completed: February 11, 2003, 18:35:19
 Job time : 30.6082 secs



DNA helicase I (EC
hypothetical prote
HIV-1 retropepsin
probable rpiR-fam
hypothetical prote
hypothetical prote
protein [imported
UL53 protein - hum
probable isomerase
copper-transportin
hypothetical prote
F6f3.12 protein -
nonstructural poly
Na+/H+ antiporter
probable bacteriop

30 42 40.0 1756 1 BVECAI
31 41.5 39.5 938 2 G70472
32 41 39.0 129 2 S63739
33 41 39.0 232 2 A10366
34 41 39.0 298 2 A85322
35 41 39.0 298 2 T05874
36 41 39.0 302 2 E95387
37 41 39.0 376 1 Q0BEW2
38 41 39.0 783 2 E91124
39 41 39.0 783 2 A85969
40 41 39.0 806 2 E83719
41 41 39.0 1034 2 S76134
42 41 39.0 1483 2 E86143
43 41 39.0 2492 1 C44213
44 40.5 38.6 490 2 F84154
45 40 38.1 115 2 AG0688

OM protein - protein search, using sw model
Run on: February 11, 2003, 18:14:11 ; Search time 21.9072 Seconds
(without alignments)
109.706 Million cell updates/sec

Title: US-09-380-738a-2
Perfect score: 105
Sequence: 1 RELEINVPGEIVEXLXXXXESITR 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues 283224
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	96.2	224	1 KBBOA2	beta-casein precu
2	96	91.4	209	2 A59068	beta-casein varian
3	91	86.7	48	2 A26334	beta-casein - wate
4	76	72.4	222	2 JC1384	beta-casein precu
5	76	72.4	222	2 A32979	beta-casein precu
6	67	63.8	232	2 A48384	beta-casein - pig
7	48	45.7	640	2 B96502	hypothetical prote
8	47	44.8	1289	2 AB2217	hypothetical prote
9	46	43.8	228	2 B70564	beta-casein precu
10	46	43.8	334	2 B84432	hypothetical prote
11	46	43.8	438	2 B72654	probable histidyl-
12	45	42.9	587	1 E69114	fumarate reductase
13	45	42.9	587	2 A90394	bps2 protein homol
14	45	42.9	807	2 T32463	hypothetical prote
15	44	41.9	218	2 H71233	hypothetical prote
16	43	41.0	98	1 B0RT2	prostatic steroid-
17	43	41.0	472	2 D71076	probable glutamate
18	43	41.0	974	2 S34189	starch phosphoryla
19	43	41.0	983	2 H72510	probable ribonucle
20	43	41.0	2492	1 MNWVTD	nonstructural poly
21	42.5	40.5	140	2 H64629	hypothetical prote
22	42	40.0	90	2 A69949	hypothetical prote
23	42	40.0	198	2 T10001	replication initia
24	42	40.0	259	2 T36003	hypothetical prote
25	42	40.0	264	2 T43997	hypothetical prote
26	42	40.0	265	1 Q0BES6	hypothetical prote
27	42	40.0	293	2 AH0633	XIRF2 protein - hu
28	42	40.0	521	2 A86909	probable transcrip
29	42	40.0	916	2 H69161	hypothetical prote DNA helicase II -

ALIGNMENTS

RESULT 1

KBBOA2

beta-casein precursor - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 24-Apr-1984 #sequence_revision 12-May-1995 #text_change 11-May-2000

C:Accession: I45873; B29087; S01860; A25846; S02429; A90489; A91191; B91192; C91192;

R:Bonsing, J.; Ring, J.M.; Stewart, A.F.; Mackinlay, A.G.

Aust. J. Biol. Sci. 41, 527-537, 1988

A:Title: Complete nucleotide sequence of the bovine beta-casein gene.

A:Reference number: I45873; MUID:90147279; PMID:3271384

A:Accession: I45873

A>Status: translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-81, 'H', '83-224 <BON>

A:Cross-references: GB:M55158; NID:gl62804; PIDN:AAA30431.1; PID:gl62805

R:Stewart, A.F.; Bonsing, J.; Beattie, C.W.; Shah, F.; Willis, I.M.; Mackinlay, A.G.

Mol. Biol. Evol. 4, 231-241, 1987

A:Title: Complete nucleotide sequences of bovine alpha-s2- and beta-casein cDNAs: c

A:Reference number: A93062; MUID:88188989; PMID:2833669

A:Accession: B29087

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-224 <SFE>

A:Cross-references: GB:M16645; NID:gl62930; PIDN:AAA30480.1; PID:gl62931

A:Experimental source: A2 variant

R:Baev, A.A.; Smirnov, I.K.; Gorodetskii, S.I.

Mol. Biol. 21, 214-222, 1987

A:Title: Primary structure of bovine beta-casein cDNA.

A:Reference number: S01860

A:Accession: S01860

A:Molecule type: mRNA

A:Residues: 1-81, 'H', '83-224 <BAE>

A:Cross-references: EMBL:X06359; NID:gl71; PIDN:CAA29658.1; PID:g757752

A:Experimental source: A1 variant

A>Note: This paper is a translation of the Russian paper published in Mol. Biol. Mo

R:Timenez-Flores, Y.C.; Kang, Y.C.; Richardson, T.

Biochem. Biophys. Res. Commun. 142, 617-621, 1987

A:Title: Cloning and sequence analysis of bovine beta-casein cDNA.

A:Reference number: A25846; MUID:87128158; PMID:3814153

A:Accession: A25846

A:Molecule type: mRNA

A:Residues: 1-107, 'L', '109-151, 'PL', '154-209, 'Q', '211-224 <JIM>

A:Cross-references: GB:M15132; NID:gl62796; PIDN:AAA30430.1; PID:gl62797

R:Carles, C.; Huet, J.C.; Ribadeau-Dumas, B.

FEBS Lett. 229, 265-272, 1988

A:Title: A new strategy for primary structure determination of proteins: applicatio

A:Reference number: S02429; MUID:88152252; PMID:3278933

A:Accession: S02429

A:Molecule type: protein

A:Residues: 16-81, 'H', '83-224 <CAR>

A:Experimental source: A1 variant

R:Yan, S.B.; Wold, F.

Biochemistry 23, 3759-3765, 1984
A:Title: Neolycoproteins: in vitro introduction of glycosyl units at glutaminesin beta-
A:Reference number: A90489; MUID:85000478; PMID:6148101
A:Accession: A90489
A:Molecule type: protein
R:Ribadeau-Dumas, B.; Brignon, G.; Grosclaude, F.; Mercier, J.C.
Eur. J. Biochem. 25, 505-514, 1972
A:Title: Structure primaire de la caseine beta bovine.
A:Reference number: A91191; MUID:72233212; PMID:4557764
A:Accession: A91191
A:Molecule type: protein
A:Residues: 16-131, 'Q', 133-151, 'PL', 154-189, 'E', 191-209, 'Q', 211-224 <RIB>
A:Experimental source: A2 variant
A:Note: article in French with an English abstract
R:Grosclaude, F.; Mahe, M.F.; Mercier, J.C.; Ribadeau-Dumas, B.
Eur. J. Biochem. 26, 328-337, 1972
A:Title: Caracterisation des variants genetiques des caseines alpha-S1 et beta bovines.
A:Reference number: A91192; MUID:72214259; PMID:5064450
A:Note: article in French with an English abstract
A:Accession: B91192
A:Molecule type: protein
A:Residues: 16-81, 'H', 83-131, 'Q', 133-151, 'PL', 154-189, 'E', 191-209, 'Q', 211-224 <VAL>
A:Experimental source: A1 variant
A:Accession: C91192
A:Molecule type: protein
A:Residues: 16-81, 'H', 83-131, 'Q', 133-136, 'R', 138-151, 'PL', 154-189, 'E', 191-209, 'Q', 211-224
A:Experimental source: B variant
A:Accession: D91192
A:Molecule type: protein
A:Residues: 16-51, 'K', 53-81, 'H', 83-131, 'Q', 133-151, 'PL', 154-189, 'E', 191-209, 'Q', 211-224
A:Experimental source: C variant
A:Note: this variant lacks a phosphate group on 50-Ser
R:Ribadeau-Dumas, B.; Grosclaude, F.; Mercier, J.C.
C. R. Acad. Sci. Hebd. Seances Acad. Sci. D 270, 2369-2372, 1970
A:Title: Localisation dans la chaine peptidique de la caseine beta bovine de la substitution
A:Reference number: A90739; MUID:71452171; PMID:4997616
A:Note: article in French with an English abstract
A:Accession: A90739
A:Molecule type: protein
A:Residues: 118-120, 'Q', 122-124 <VA3>
A:Experimental source: A3 variant
R:Simons, G.; van den Heuvel, W.; Reynen, T.; Frijters, A.; Rutten, G.; Slangen, C.J.; G
Protein Eng. 6, 763-770, 1993
A:Title: Overproduction of bovine beta-casein in Escherichia coli and engineering of its
A:Reference number: I46963; MUID:94068382; PMID:8248100
A:Accession: I46963
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-120, 'Q', 122-224 <SIM>
A:Cross-references: GB:S67277; NID:G459291; PIDN:AAB29137.1; PID:G459292
A:Experimental source: A3 variant
R:Grosclaude, F.; Mahe, M.F.; Voglino, G.F.
FEBS Lett. 45, 3-5, 1974
A:Title: Le variant beta-E et le code de phosphorylation des caseines bovines.
A:Reference number: A91413; MUID:75005247; PMID:4411121
A:Note: article in French with an English abstract
A:Accession: A91413
A:Molecule type: protein
A:Residues: 48-50, 'K', 52-63 <VAE>
A:Experimental source: E variant
A:Note: 50-Ser is phosphorylated
C:Comment: The sequence shown is the A2 variant.
C:Genetics:
A:Introns: 17/3; 26/3; 35/3; 43/3; 57/3; 223/3
C:Superfamily: beta-casein
C:Keywords: milk; phosphoprotein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-224/Product: beta-casein #status experimental <WAT>
F:30-32, 33, 34/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status exp
F:50/Binding site: phosphate (Ser) (covalent) (by casein kinase II) (partial) #status ex
Query Match 96.2%; Score 101; DB 1; Length 224;

Best Local Similarity 84.0%; Pred. No. 6.4e-09;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 RELEELNVPGIVEIXLXXEESITR 25
|||||
Db 16 RELEELNVPGIVEISLSSEESITR 40
|||||
RESULT 2
A59068
beta-casein variant CnH - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Sep-1999 #sequence_revision 24-Sep-1999 #text_change 24-Sep-1999
C:Accession: A59068; B59068
R:Han, S.K.; Shin, Y.C.
Anim. Genet. 27(Suppl.2), 91b, 1996
A:Title: Biochemical characterization of the new beta-casein variant in Korean cattl
A:Reference number: A59068
A:Accession: A59068
A>Status: protein sequence not shown
A:Molecule type: protein
A:Residues: 1-209 <HAN1>
A:Experimental source: strain Korean cattle
A:Note: submitted to the Protein Sequence Database, September 1999
A:Note: includes casein phosphopeptide H
A:Accession: B59068
A>Status: protein sequence not shown
A:Molecule type: protein
A:Residues: 1-28 <HAN2>
A:Experimental source: strain Korean cattle
C:Superfamily: beta-casein
C:Keywords: milk; phosphoprotein
F:15,17,18,19/Binding site: phosphate (Ser) (covalent) #status predicted
Query Match 91.4%; Score 96; DB 2; Length 209;
Best Local Similarity 83.3%; Pred. No. 3.9e-08;
Matches 20; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 RELEELNVPGIVEIXLXXEESIT 24
|||||
Db 1 RELEELNVPGIVEISLSSEESIT 24
|||||
RESULT 3
A26334
beta-casein - water buffalo (fragment)
C:Species: Bubalus arnee (water buffalo)
C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 30-Sep-1993
C:Accession: A26334
R:Petrilli, P.; Pucci, P.; Morris, H.R.; Addeo, F.
Biochem. Biophys. Res. Commun. 140, 28-37, 1986
A:Title: Assignment of phosphorylation sites in buffalo beta-casein by fast atom bomb
A:Reference number: A26334; MUID:87048757; PMID:3778448
A:Accession: A26334
A:Molecule type: protein
A:Residues: 1-48 <PET>
C:Superfamily: beta-casein
Query Match 86.7%; Score 91; DB 2; Length 48;
Best Local Similarity 82.6%; Pred. No. 4.8e-08;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 RELEELNVPGIVEIXLXXEESIT 24
|||||
Db 2 RELEELNVPGIVEISLSSEESIT 24
|||||
RESULT 4
JC1384
beta-casein precursor - goat
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 10-Jun-1993 #sequence_revision 10-Jun-1993 #text_change 23-Feb-1997
C:Accession: JC1384

us-09-380-738a-2.rpr

Wed Feb 12 11:35:22 2003

R:Roberts, B.; DiTullio, P.; Vitale, J.; Hehir, K.; Gordon, K.
Gene 121, 255-262, 1992

A:Title: Cloning of the goat beta-casein-encoding gene and expression in transgenic mice
A:Reference number: JCI384; MUID:93077039; PMID:1446822

A:Accession: JCI384

A:Molecule type: DNA

A:Residues: 1-222 <ROB>

A:Cross-references: GB:M90556

C:Genetics:

A:Gene: CSN2

A:Introns: 17/3; 26/3; 35/3; 45/3; 57/3; 221/3

C:Superfamily: beta-casein

C:Keywords: milk; phosphoprotein

Query Match 72.4%; Score 76; DB 2; Length 222;
Best Local Similarity 70.8%; Pred. No. 7; 7e-05;
Matches 17; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 RELEELNVPGEIVEXLXXEESIT 24

DB 16 RQEEELNVGGEIVESLSSEESIT 39

RESULT 5

A32979

beta-casein precursor - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C>Date: 12-Oct-1989 #sequence_revision 31-Dec-1993 #text_change 13-Aug-1999

C:Accession: A32979; A29173

R:Provot, C.; Persuy, M.A.; Mercier, J.C.

Biochimie 71, 827-832, 1989

A:Title: Complete nucleotide sequence of ovine beta-casein cDNA: inter-species comparison

A:Reference number: A32979; MUID:89375530; PMID:2505862

A:Accession: A32979

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-222 <PRO>

A:Cross-references: GB:X16482; NID:g1210; PIDN:CAA34502.1; PID:g1211

A>Note: The authors translated the codon CAC for residue 160 as Lys

R:Richardson, B.C.; Mercier, J.C.

Eur. J. Biochem. 99, 285-297, 1979

A:Title: The primary structure of the ovine beta-caseins.

A:Reference number: A29173; MUID:80046695; PMID:499202

A:Accession: A29173

A:Molecule type: protein

A:Residues: 16-69, 'T', '71-77', 'P', '79-81', 'A', '83-222 <RIC>

C:Superfamily: beta-casein

C:Keywords: milk; phosphoprotein

F:1-15/Domain: signal sequence #status predicted <STG>

F:16-222/Product: beta-casein #status experimental <MAT>

F:30,32,33,34,50/Binding site: phosphate (Ser) (covalent) (by casein kinase II) #status

Query Match 72.4%; Score 76; DB 2; Length 222;

Best Local Similarity 70.8%; Pred. No. 7; 7e-05;

Matches 17; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 RELEELNVPGEIVEXLXXEESIT 24

DB 16 RQEEELNVGGEIVESLSSEESIT 39

RESULT 6

A48384

beta-casein - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 03-May-1996

C:Accession: A48384

R:Alexander, L.J.; Beattie, C.W.

Anim. Genet. 23, 369-371, 1992

A:Title: The sequence of porcine beta-casein cDNA.

A:Reference number: A48384; MUID:92367961; PMID:1503277

A:Accession: A48384

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-232 <ALE>

A:Experimental source: mammary gland

A>Note: sequence inconsistent with the nucleotide translation

C:Superfamily: beta-casein

Query Match 63.8%; Score 67; DB 2; Length 232;

Best Local Similarity 62.5%; Pred. No. 0.0024;

Matches 15; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 RELEELNVPGEIVEXLXXEESIT 24

DB 16 RAKEELNASGETVESLSSEESIT 39

RESULT 7

B96502

hypothetical protein F28H19.8 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: B96502

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Haizhar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talic

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: B96502

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-640 <STO>

A:Cross-references: GB:AE005173; NID:g7523674; PIDN:AAF63114.1; GSPDB:GN00141

C:Genetics:

A:Gene: F28H19.8

A:Map position: 1

Query Match 45.7%; Score 48; DB 2; Length 640;

Best Local Similarity 50.0%; Pred. No. 9.6;

Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 EELNVPGEIVEXLXXEESITR 25

DB 70 BELNICEEVSSVTKEESSTR 91

RESULT 8

AB2217

hypothetical protein all3289 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AB2217

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iri

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Taba

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacteriu

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AB2217

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1289 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAB74988.1; PID:g17132384; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all3289

Query Match 44.8%; Score 47; DB 2; Length 1289;

1 RELEELNPGEIVEXLXXX--EESITR 25

2 protein homolog (bps2) [imported] - Sulfolobus solfataricus
Species: Sulfolobus solfataricus
Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 2

A:Gene: PH0130
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH0130
Query Match 41.9%; Score 44; DB 2; Length 218;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
QY 1 RELEELNVPGEIVEEXLXXE 20
Db 140 QNLNELNLPGEIRLLIVPE 159
Search completed: February 11, 2003, 18:21:14
Job time : 23.9072 secs

C:Accession: A90394
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awauez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, H.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: A90394
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-587 <KUR>
A:Cross-references: GB:AE006641; NID:g13815540; PIDN:AAK42408.1; GSPDB:GN00155
C:Genetics:
A:Gene: bps2

Query Match 42.9%; Score 45; DB 2; Length 587;
Best Local Similarity 45.5%; Pred. NO. 27;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 EELELNPGEIVEEXLXXEESI 23
Db 425 ELQLGIPSSILEELKEKEHI 446

RESULT 14
T32463
hypothetical protein F52G3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
C:Accession: T32463
R:Blanchard, M.; Gattung, S.; Sansone, J.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid F52G3.
A:Reference number: Z21173
A:Accession: T32463
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-807 <BLA>
A:Cross-references: EMBL:AF026212; PIDN:AAB71298.1; GSPDB:GN00028; CESP:F52G3.3
A:Experimental source: strain Bristol N2; clone F52G3
C:Genetics:
A:Gene: CESP:F52G3.3
A:Map position: X
A:Introns: 383/3; 434/2; 507/3; 550/3; 648/3; 759/2

Query Match 42.9%; Score 45; DB 2; Length 807;
Best Local Similarity 45.0%; Pred. NO. 39;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 EELELNPGEIVEEXLXXEE 21
Db 277 ELDELVDGNGVDPGLTAE 296

RESULT 15
H71233
hypothetical protein PH0130 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 28-Jul-2000
C:Accession: H71233
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki-
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: H71233
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-218 <KAW>
A:Cross-references: GB:AP000001; NID:g3236128; PIDN:BAA29199.1; PID:g3256516
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:07:35 ; Search time 11.0825 Seconds
(without alignments)
93.563 Million cell updates/sec

Title: US-09-380-738a-2

Perfect score: 105

Sequence: 1 RELELNVPGEIVEXLXXXESYIR 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	101	96.2	224	1 CASB_BOVIN	P02666 bos taurus
2	96	91.4	224	1 CASB_BUBBU	Q9tsi0 bubalus bub
3	76	72.4	222	1 CASB_CAPHI	P33048 capra hircu
4	76	72.4	222	1 CASB_SHEEP	P11839 ovis aries
5	67	63.8	232	1 CASB_PIG	P39037 sus scrofa
6	58	55.2	232	1 CASB_CAMDR	Q9tyd0 camelus dro
7	46	43.8	228	1 CASB_RABIT	Q9yeb2 aeropyrum p
8	46	43.8	438	1 SVH_AERPE	P14565 escherichia
9	44	41.9	1756	1 TRI1_ECOLI	P02781 rattus norv
10	43	41.0	112	1 PSC2_RAT	P53535 solanum tub
11	43	41.0	974	1 PHS2_SOLTU	P27282 venezuelan
12	43	41.0	2492	1 POLN_EEVT	P45937 bacillus su
13	42	40.0	90	1 YOQB_BACSU	P28865 human herpe
14	42	40.0	264	1 UL31_HSV6U	P46388 mycobacteri
15	42	40.0	502	1 DNAA_MYCLE	P22706 escherichia
16	42	40.0	1756	1 TRI2_ECOLI	Q9fnp7 arabidopsis
17	41	39.0	178	1 AG31_ARATH	Q9t091 arabidopsis
18	41	39.0	298	1 VP26_ARATH	P16794 human cytom
19	41	39.0	376	1 UL53_HCMVA	P41725 felis silve
20	41	39.0	417	1 PGCB_FELCA	P74299 synchocyst
21	41	39.0	1034	1 CAPP_SYNY3	P29813 e genome po
22	41	39.0	2195	1 POLG_ECL1G	P36327 venezuelan
23	41	39.0	2485	1 POLN_EEUV3	P05814 homo sapien
24	40	38.1	226	1 CASB_HUMAN	P55178 staphylococ
25	40	38.1	234	1 YAG5_STALU	P52361 human herpe
26	40	38.1	259	1 UL31_HSV7J	O67891 aquifex aeo
27	40	38.1	282	1 PANC_AQUAE	P09534 xenopus lae
28	40	38.1	360	1 DVRI_XENLA	P09534 xenopus lae
29	40	38.1	365	1 RNN9_YEAST	P53437 saccharomyc
30	40	38.1	397	1 HMDH_METTH	O26662 methanobact
31	40	38.1	455	1 VNS1_BMDNV	P05840 bombyx dens
32	40	38.1	862	1 LOX1_HORVU	P29114 hordeum vul
33	40	38.1	880	1 RA50_PYPAB	Q9uzc8 pyrococcus

34 39 37.1 35 1 PBP_HYACE
35 39 37.1 109 1 VGLI_HSVSB
36 39 37.1 226 1 PDGA_XENLA
37 39 37.1 318 1 Y692_METTH
38 39 37.1 368 1 Y1U5_YEAST
39 39 37.1 501 1 YM05_YEAST
40 39 37.1 579 1 UL25_VZVD
41 39 37.1 675 1 ATKB_DEIRA
42 39 37.1 803 1 ATCU_BACSU
43 39 37.1 825 1 GUN3_BACSA
44 39 37.1 1057 1 CARB_STAAM
45 39 37.1 1057 1 CARB_STAAM

ALIGNMENTS

RESULT 1
CASB_BOVIN STANDARD; PRT; 224 AA.
ID CASB_BOVIN STANDARD; PRT; 224 AA.
AC P02666;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta casein precursor.
GN CSN2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID:9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Baev A.A., Smirnov I.K., Gorodetsky S.I.;
RT "Primary structure of bovine beta-casein cDNA."
RL Mol. Biol. (Mosk) 21:214-222(1987).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=88188989; PubMed=2833669;
RT Stewart A.F., Bonsing J., Beattie C.W., Shah F., Willis I.M.,
RA Mackinlay A.G.;
RT "Complete nucleotide sequences of bovine alpha S2- and beta-casein
cDNAs; comparisons with related sequences in other species."
RL Mol. Biol. Evol. 4:231-241(1987).
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=90147279; PubMed=3271384;
RT Bonsing J., Ring J.M., Stewart A.F., Mackinlay A.G.;
RA "Complete nucleotide sequence of the bovine beta-casein gene."
RT Aust. J. Biol. Sci. 41:527-537(1988).
RN [4]
RP SEQUENCE FROM N.A.
RA MEDLINE=87128158; PubMed=3814153;
RT Jimenez-Flores R., Kang Y.C., Richardson T.;
RA "Cloning and sequence analysis of bovine beta-casein cDNA."
RT Biochem. Biophys. Res. Commun. 142:617-621(1987).
RN [5]
RP SEQUENCE FROM N.A. (VARIANT A3).
RA TISSUE=Mammary gland;
RC MEDLINE=94068382; PubMed=8248100;
RX Simons G., van den Heuvel W., Reynen T., Frijters A., Rutten G.,
RA Slangen C.J., Groenen M., de Vos W.M., Siezen R.J.;
RT "Overproduction of bovine beta-casein in Escherichia coli and
engineering of its main chymosin cleavage site."
RL Protein Eng. 6:763-770(1993).
RN [6]
RP SEQUENCE OF 16-224 (VARIANT A2).
RA MEDLINE=8815222; PubMed=3278933;
RT Charles C., Huet J.-C., Ribadeau-Dumas B.;
RA "A new strategy for primary structure determination of proteins:
application to bovine beta-casein."
RL FEBS Lett. 229:265-272(1988).
RN [7]

wed Feb 12 11:35:22 2003

Db 16 RELEELNVPGEIVESLSSESIT 39

RESULT 3

QY 1 RELEELNVPGEIVEXLXXEESITR 25

DB 16 RELEELNVPGEIVESLSSESITR 40

Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

RESULT 2

ID CASEB_BUBBU STANDARD; PRT; 224 AA.

AC Q9TS10; O62824;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Beta casein precursor.

GN CSN2.

OS Bubalus bubalis (Domestic water buffalo).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bubalus.

OX NCBI_TaxID=89462;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Mammary gland;

RA Klotz A., Krause I., Einspanier R.;

RT "Isolation of mRNA from buffalo milk.";

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Mammary gland;

RA Das P., Garg L.C.;

RT "cDNA cloning and sequencing of beta-casein gene in B. bubalis.";

RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: IMPORTANT ROLE IN DETERMINATION OF THE SURFACE

CC PROPERTIES OF THE CASEIN MICELLES.

CC -1- SUBCELLULAR LOCATION: Extracellular.

CC -1- TISSUE SPECIFICITY: Mammary gland; MILK.

CC -1- SIMILARITY: BELONGS TO THE BETA-CASEIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; M90561; AAA30906.1; -

DR EMBL; M90556; AAA30906.1; JOINED.

DR EMBL; M90557; AAA30906.1; JOINED.

DR EMBL; M90558; AAA30906.1; JOINED.

DR EMBL; M90560; AAA30906.1; JOINED.

DR PIR; JCI384; JCI384.

DR InterPro; IPR001588; Casein.

DR Pfam; PF00363; caseins; 1.

DR PROSITE; PS00306; CASEIN_ALPHA_BETA; FALSE_NEG.

DR Milk; Phosphorylation; Glycoprotein; Signal.

KW SIGNAL

FT CHAIN

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

FT MOD_RES

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; X13043; CAA31449.1; -
EMBL; M33582; AAA31168.1; -
PIR; J00564; J00564.
InterPro; IPR001588; Casein.
DR Pfam; PF00363; caseins; 1.
DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
KW Milk; Phosphorylation; Glycoprotein; Signal.
FT SIGNAL 1 15
FT CHAIN 16 228 BETA CASEIN.
FT MOD_RES 30 30 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 32 32 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 33 33 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 34 34 PHOSPHORYLATION (POTENTIAL).
FT CONFLICT 87 87 I -> T (IN REF. 2).
SQ SEQUENCE 228 AA; 26072 MW; 16BB2D7C78887E8A CRC64;

Query Match 43.8%; Score 46; DB 1; Length 228;
Best Local Similarity 47.8%; Pred. No. 2.9;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 RELEENVPGEIVEXLXXEESIT 21
DB 16 REKEEFTAGEALESTSSSEESIT 36

RESULT 8

SYN_AERPE SYH_AERPE STANDARD; PRT; 438 AA.
AC QYEB2;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine-tRNA ligase) (HISRS).
DE HISS OR APE0662.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kwarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y., Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H., Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K., Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RA "Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, Aeropyrum pernix K1.";
RT DNA Res. 6:83-101(1999).
RL -!- CATALYTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP + diphosphate + L-histidyl-tRNA(His).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-tRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AP000060; BAA79634.1; -
HSSP; 032422; 10E0.

-!- FUNCTION: IMPORTANT ROLE IN DETERMINATION OF THE SURFACE PROPERTIES OF THE CASEIN MICELLES (BY SIMILARITY).
-!- SUBCELLULAR LOCATION: Extracellular.
-!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
-!- SIMILARITY: BELONGS TO THE BETA-CASEIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AJ012630; CAA10079.1; -
InterPro; IPR001588; Casein.
DR Pfam; PF00363; caseins; 1.
DR PROSITE; PS00306; CASEIN_ALPHA_BETA; FALSE_NEG.
KW Milk; Phosphorylation; Glycoprotein; Signal.
FT SIGNAL 1 15
FT CHAIN 16 232 BETA CASEIN.
FT MOD_RES 30 30 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 32 32 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 33 33 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 34 34 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 232 AA; 26218 MW; A0F9F41D2EA7C518 CRC64;

Query Match 55.2%; Score 58; DB 1; Length 232;
Best Local Similarity 50.0%; Pred. No. 0.034;
Matches 12; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 RELEENVPGEIVEXLXXEESIT 24
DB 16 REKEEFTAGEALESTSSSEESIT 39

RESULT 7

CASB_RABIT CASB_RABIT STANDARD; PRT; 228 AA.
AC P09116;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Beta casein precursor.
GN CSN2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Mammary gland;
RX MEDLINE=89098330; PubMed=3211753;
RA Devino E., Schaefer E., Jollivet G., Fontaine M.L., Kraehenbuhl J.P., Houdebine L.M.;
RA "Sequence of the rabbit beta-casein cDNA: comparison with other casein cDNA sequences.";
RT Nucleic Acids Res. 16:11814-11814(1988).
RL [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91153663; PubMed=1999295;
RA Thepot D., Devino E., Fontaine M.L., Houdebine L.M.;
RA "Structure of the gene encoding rabbit beta-casein.";
RL Gene 97:301-306(1991).
CC -!- FUNCTION: IMPORTANT ROLE IN DETERMINATION OF THE SURFACE PROPERTIES OF THE CASEIN MICELLES.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: BELONGS TO THE BETA-CASEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

DR InterPro: IPR002106; AAcRNA_ligaseII.
 DR InterPro: IPR004154; HGTP_anticoodon.
 DR InterPro: IPR004516; HisS.
 DR InterPro: IPR002314; tRNA-synt_2b.
 DR Pfam: PF00587; tRNA-synt_2b; 1.
 DR Pfam: PF03129; HGTP_anticoodon; 1.
 DR TIGRFAMs: TIGR00442; HisS; 1.
 DR PROSITE: PS00862; AA_TRNA_LIGASE_II; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome.
 SQ SEQUENCE 438 AA: 49111 MW; ECCE67F59A9FB7BC CRC64;

Query Match 43.88; Score 46; DB 1; Length 438;
 Best Local Similarity 56.28; Pred. No. 5.9;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 RELEELNVPGEIVEXL 16
 ||||| : || :
 Db 205 RELERLGLPGSTVERI 220

RESULT 9

TRIL_ECOLI STANDARD; PRT; 1756 AA.
 ID TRIL_ECOLI
 AC P14565; Q51811;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Trai protein (DNA helicase I) (EC 3.6.1.-) [Contains: Trai* protein].
 GN TRAI
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=90299847; PubMed=2163400;
 RA Bradshaw H.D. Jr., Traxler B.A., Minkley E.G. Jr., Nester E.W.,
 RA Gordon M.P.;
 RT "Nucleotide sequence of the traI (helicase I) gene from the sex
 factor F.,"
 RL J. Bacteriol. 172:4127-4131(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94359430; PubMed=7915817;
 RA Frost L.S., Ippen-Ihler K., Skurray R.A.;
 RT "Analysis of the sequence and gene products of the transfer region of
 the F sex factor.,"
 RL Microbiol. Rev. 58:162-210(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN-K12 / CR63;
 RA Shimizu H., Saitoh Y., Suda Y., Uehara K., Sampei G., Mizobuchi K.;
 RT "Complete nucleotide sequence of the F plasmid: its implications for
 organization and diversification of plasmid genomes.,"
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-150 FROM N.A.
 RX STRAIN-K12;
 RA MEDLINE=90034191; PubMed=2680768;
 RA Jajajakumari M.B., Manning P.A.;
 RT "Nucleotide sequence of the traD region in the Escherichia coli F sex
 factor.,"
 RL Gene 81:195-202(1989).
 RN [5]
 RP SEQUENCE OF 1-72 FROM N.A.
 RX MEDLINE=90317835; PubMed=2164585;
 RA Yoshioke Y., Fujita Y., Ohtsubo E.;
 RT "Nucleotide sequence of the promoter-distal region of the tra operon
 of plasmid R100, including traI (DNA helicase I) and traD genes.,"
 RL J. Mol. Biol. 214:39-53(1990).
 RN [6]

RP SEQUENCE OF 955-1756 FROM N.A.
 RX MEDLINE=96347127; PubMed=8736534;
 RA Penfold S.S., Simon J., Frost L.S.;
 RT "Regulation of the expression of the traM gene of the F sex factor of
 Escherichia coli.,"
 RL Mol. Microbiol. 20:549-558(1996).
 CC -!- FUNCTION: TRAI HAS BEEN IDENTIFIED AS DNA HELICASE I AND IT
 CC ALSO HAS AN ADDITIONAL ACTIVITY OF SITE-SPECIFIC NICKING AT
 CC ORIT. DNA HELICASE I IS A POTENT DNA-DEPENDENT ATPASE.
 CC -!- ALTERNATIVE PRODUCTS: THE TRAI* PROTEIN ARISES FROM AN INTERNAL
 CC IN-FRAME TRANSLATIONAL START IN TRAI.
 CC -!- SIMILARITY: STRONG TO TRAI OF PLASMID INCFII R100.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M54796; AAA98085.1; -;
 DR EMBL; M54796; AAA98086.1; -;
 DR EMBL; U01159; AAC44186.1; -;
 DR EMBL; AP001918; BAA97974.1; -;
 DR EMBL; M29254; AAA83930.1; ALT_INIT.
 DR EMBL; X57430; CAA40677.1; -;
 DR EMBL; U01159; AAC44187.1; -;
 DR PIR; PS0068; BVECAL.
 DR EcoGene; EG40119; traI.
 KW Plasmid; Helicase; ATP-binding; DNA-binding; Conjugation; Hydrolase;
 KW Alternative initiation; Complete proteome.
 FT CHAIN 1 1756
 FT TRAI PROTEIN.
 FT INIT_MET 955 1756
 FT FOR TRAI*.
 FT NP_BIND 992 999
 FT ATP (POTENTIAL).
 FT CONFLICT 69 74 MODGSN -> CRMAVT (IN REF. 4).
 SQ SEQUENCE 1756 AA; 192015 MW; AA07D61DB2BFD9FA CRC64;

Query Match 41.98; Score 44; DB 1; Length 1756;
 Best Local Similarity 44.48; Pred. No. 58;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 8 VPGGEIVEXLXXXEESITR 25
 ||||| : : ||
 Db 961 VPGELMEKLTSGRAATR 978

RESULT 10

PSC2_RAT STANDARD; PRT; 112 AA.
 ID PSC2_RAT
 AC P02781;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Prostetic steroid-binding protein C2 chain precursor (Prostatein
 DE peptide C2).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87146484; PubMed=2881277;
 RA Delaey B., Dirckx L., Decourt J.-L., Claessens F., Peeters B.,
 RA Rombauts W.;
 RT "Rat prostetic binding protein: the complete sequence of the C2 gene
 RT and its flanking regions.,"
 RL Nucleic Acids Res. 15:1627-1641(1987).
 RN [2]
 RP SEQUENCE OF 21-112.
 RX MEDLINE=83209619; PubMed=6343081;
 RA Peeters B., Heyns W., Mous J., Rombauts W.;

Wed Feb 12 11:35:22 2003

RT "Structural studies on rat prostatic binding protein. The primary
 RT structure of component C2 from subunit S.";
 RL Eur. J. Biochem. 132:669-679(1983).
 RN [3]
 RN SEQUENCE OF 1-100 FROM N.A.
 RP MEDLINE-8220075; PubMed-6896362;
 RX Parker M., Needham M., White R.;
 RA "Prostatic steroid binding protein: gene duplication and steroid
 RT binding.";
 RL Nature 298:92-94(1982).
 CC -!- FUNCTION: PART OF PROSTATEIN WHICH IS THE MAJOR SECRETORY
 CC GLYCOPROTEIN OF VENTRAL PROSTATE GLAND.
 CC -!- SUBUNIT: PROSTATEIN IS COMPOSED OF THREE DIFFERENT PEPTIDES CALLED
 CC C1, C2 AND C3. THESE FORM COVALENT C1-C3 (F) AND C2-C3 (S)
 CC HETERODIMERS, THESE NONCOVALENT ASSOCIATION FORMS TETRAMERIC
 CC (C1-C3/C3-C2) PROSTATEIN MOLECULES.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PM: LINKED BY THREE DISULFIDE BONDS TO C3.
 CC -!- MISCELLANEOUS: THE HETERODIMER CAN BIND NON-POLAR STEROIDS,
 CC CHOLESTEROL AND A GROUP OF SMALL, PROLINE-RICH PEPTIDES.
 CC -!- SIMILARITY: BELONGS TO THE UTEROGLOBIN FAMILY. LIPOPHILIN
 CC SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X05034; CAA28708.1; -;
 CC EMBL; V01256; CAA24569.1; -;
 CC EMBL; J00776; AAA51641.1; -;
 CC PIR; A03251; B0RT2.
 CC PIR; A26671; A26671.
 CC InterPro: IPR000329; Uteroglobin.
 CC PROSITE; PS00403; UTEROGLOBIN_1; FALSE_NEG.
 CC PROSITE; PS00404; UTEROGLOBIN_2; FALSE_NEG.
 CC Steroid-binding; Signal.
 CC SIGNAL 1 20
 CC CHAIN 21 112 PROSTATIC STEROID-BINDING PROTEIN C2
 CC CHAIN.
 CC MOD_RES 21 21 BLOCKED.
 CC FT DISULFID 28 28 INTERCHAIN (WITH C3) (PROBABLE).
 CC FT DISULFID 69 69 INTERCHAIN (WITH C3) (PROBABLE).
 CC FT DISULFID 92 92 INTERCHAIN (WITH C3) (PROBABLE).
 CC FT DISULFID 26 26 MISSING (IN REF. 3).
 CC FT CONFLICT 88 88 I -> T (IN REF. 3).
 CC FT CONFLICT 96 112 VMLQINPGRWFSEIN -> YGYK (IN REF. 3).
 CC FT CONFLICT 96 112
 CC SEQUENCE 112 AA; 12828 MW; DA65A6A8E677864 CRC64;
 CC
 CC Query Match 41.0%; Score 43; DB 1; Length 112;
 CC Best Local Similarity 64.3%; Pred. No. 3.9;
 CC Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 1 RELEELNVPGEIVE 14
 CC 48 RELEEFADPEAVE 61
 CC
 CC RESULT 11
 CC PHS2_SOLTU STANDARD; PRT; 974 AA.
 CC ID PHS2_SOLTU
 CC AC P53535;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Alpha-1,4 glucan phosphorylase, L-2 isozyme, chloroplast precursor
 CC DE (EC 2.4.1.1) (Starch phosphorylase L-2).
 CC GN STP-1.
 CC OS Solanum tuberosum (Potato).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. Desiree; TISSUE=Leaf;
 RX MEDLINE-95201249; PubMed-7894019;
 RA Sonnewald U., Basner A., Greve B., Steup M.;
 RT "A second L-type isozyme of potato glucan phosphorylase: cloning,
 RT antisense inhibition and expression analysis.";
 RL Plant Mol. Biol. 27:567-576(1995).
 CC -!- FUNCTION: PHOSPHORYLASE IS AN IMPORTANT ALLOSTERIC ENZYME IN
 CC CARBOHYDRATE METABOLISM. ENZYMES FROM DIFFERENT SOURCES DIFFER IN
 CC THEIR REGULATORY MECHANISMS AND IN THEIR NATURAL SUBSTRATES.
 CC HOWEVER, ALL KNOWN PHOSPHORYLASES SHARE CATALYTIC AND STRUCTURAL
 CC PROPERTIES.
 CC -!- CATALYTIC ACTIVITY: [(1,4)-alpha-D-glucosyl](N) + phosphate =
 CC [(1,4)-alpha-D-glucosyl](N-1) + alpha-D-glucose 1-phosphate.
 CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -!- SUBCELLULAR LOCATION: Chloroplast; amyloplast.
 CC -!- TISSUE SPECIFICITY: LEAF.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOGEN PHOSPHORYLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X73684; CAA52036.1; -;
 CC HSSP; P06738; IYGP.
 CC InterPro: IPR000811; GT_35.
 CC PIR; A03251; B0RT2.
 CC PROSITE; PS00102; PHOSPHORYLASE; 1.
 CC TRANSFERASE; Glycosyltransferase; Carbohydrate metabolism;
 CC Allosteric enzyme; Pyridoxal phosphate; Transit peptide; Chloroplast;
 CC Amyloplast; Multigene family.
 CC TRANSIT 1 81
 CC CHAIN 82 974
 CC CHAIN
 CC BINDING 820 820 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC SEQUENCE 974 AA; 110700 MW; 5EF8A23C237463D8 CRC64;
 CC
 CC Query Match 41.0%; Score 43; DB 1; Length 974;
 CC Best Local Similarity 36.4%; Pred. No. 44;
 CC Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
 CC
 CC QY 1 RELEELNVPGEIVEIXLXXEES 22
 CC 486 RILDNVEIPSSVLELLIKAES 507
 CC
 CC RESULT 12
 CC POLN_EEVT STANDARD; PRT; 2492 AA.
 CC ID POLN_EEVT
 CC AC P27282;
 CC DT 01-AUG-1992 (Rel. 23, Created)
 CC DT 01-AUG-1992 (Rel. 23, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Nonstructural polyprotein [Contains: Nonstructural protein NSP1;
 CC Nonstructural protein NSP2; Nonstructural protein NSP3; Nonstructural
 CC protein NSP4].
 CC DE Venezuelan equine encephalitis virus (strain Trinidad donkey).
 CC OS Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 CC OC Alphavirus.
 CC OX NCBI_TaxID=11038;
 CC RN [1]
 CC RN SEQUENCE FROM N.A.
 CC RP MEDLINE-69243175; PubMed-2524126;
 CC Kinney R.M., Johnson B.J.B., Welch J.B., Tsuchiya K.R., Trent D.W.;
 CC "The full-length nucleotide sequences of the virulent Trinidad donkey

Wed Feb 12 11:35:22 2003

```

AC P28865;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein U37.
GN U37 OR X1RF2.
OS Human herpesvirus (type 6 / strain Uganda-1102) (HHV6).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
ON NCBI_TaxID=10370;
RX MEDLINE=95266321; PubMed=7747482;
RA Gomels U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J.,
RA Martin M.E., Efsthaliou S., Craxton M., Macaulay H.A.;
RT "The DNA sequence of human herpesvirus-6: structure, coding content,
RT and genome evolution.";
RL Virology 209:29-51(1995).
[1]
[2]
SEQUENCE FROM N.A.
RP MEDLINE=91333007; PubMed=1651403;
RX tao I.A., Griffin B.E., Jones M.D.;
RA "Characterization of the DNA polymerase gene of human herpesvirus 6.";
RL J. Virol. 65:4670-4680(1991).
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL31,
CC EBV-1 29, EBV BFLF2, HCMV UL53, AND VZV 27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63804; AAA74630.1; -
CC EMBL; X83413; CAA58417.1; -
CC PIR; A40898; QOBE6S.
CC InterPro; IPR003868; Herpes_UL31-like.
CC Pfam; PF02718; Herpes_UL31; 1.
CC SEQUENCE 264 AA; 30845 MW; 5A0D8D66F01AAF94 CRC64;
SQ
Query Match 40.0%; Score 42; DB 1; Length 264;
Best Local Similarity 43.8%; Pred. No. 15;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
QY 1 RELEELNVPGEIVEIXL 16
Db 238 KLEEMDIPNEISDRL 253
:::|||||:::|
[1]
RESULT 15
DAAA_MYCLE STANDARD; PRT; 502 AA.
AC P46388;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chromosomal replication initiator protein dnaA.
GN DAAA OR ML0001.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteriales (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
ON NCBI_TaxID=1769;
RX MEDLINE=97124199; PubMed=8969512;
RA Fsihi H., de Rossi E., Salazar L., Cantoni R., Labo M., Riccardi G.,
RA Takkif H.E., Eiglmeyer K., Bergh S., Cole S.T.;
RT "Gene arrangement and organization in an approximately 76 kb fragment
RT encompassing the oric region of the chromosome of Mycobacterium
RT leprae.";
RL Microbiology 142:3147-3161(1996).
[2]
SEQUENCE FROM N.A.
RP MEDLINE=21128732; PubMed=11234002;
RX Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RA "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
RL -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
CC OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
CC IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
CC (DAAA BOX): 5'-TTATC(C/A)A(C/A)A-3'. DAAA BINDS TO ATP AND TO
CC ACIDIC PHOSPHOLIPIDS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE DAAA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L39923; AAB53141.1; ALT_INIT.
CC EMBL; AL583917; CAC29509.1; ALT_INIT.
CC Leproma; ML0001; -
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001957; Bac_DnaA.
CC Pfam; PF00308; bac.dnaA; 1.
CC PRINTS; PR00051; DAAA.
CC SMART; SM00382; DnaA; 1.
CC TIGRFAMS; TIGR00362; DnaA; 1.
CC PROSITE; PS01008; DAAA; 1.
CC DnaA replication; DnaA-binding; ATP-binding; Complete proteome.
KW NP_BIND 203 210 ATP (POTENTIAL).
FT CONFLICT 183 183 A -> R (IN REF. 1).
ET SEQUENCE 502 AA; 56313 MW; 2D45721D99DEDEAE CRC64;
SQ
Query Match 40.0%; Score 42; DB 1; Length 502;
Best Local Similarity 35.0%; Pred. No. 30;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
QY 2 ELEELNVPGEIVEIXLXXXEE 21
Db 342 QMERLAVPGDVLELIASSIE 361
:::|:|:|:|:|:|:|
Search completed: February 11, 2003, 18:17:45
Job time : 13.0825 secs

```



```

AC 09N2G8:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Beta-casein.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20541290; PubMed=11092743;
RA Watanabe M., Sugano S., Togashi T., Inai J., Uchida K., Yamaguchi R.,
RA Tateyama S.;
RT "Molecular cloning and phylogenetic analysis of canine beta-casein.";
RL DNA Seq. 11:295-300(2000).
DR EMBL: AB035080; BAA95931.1;
DR InterPro: IPR001588; Casein.
DR Pfam: PF00363; caseins; 1.
DR PROSITE: PS00306; CASEIN_ALPHA_BETA; 1.
SQ SEQUENCE 250 AA; 28401 MW; 1D58391E7BF97ED8 CRC64;

Query Match
Best Local Similarity 59.0%; Score 62; DB 6; Length 250;
Matches 14; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 1 REELNVPGEIVEXLXXEERSIT 24
Db 16 REKEELTLNETVESLSSESIT 39

RESULT 3
Q9GKK3
ID Q9GKK3 PRELIMINARY; PRT; 233 AA.
AC Q9GKK3;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE Beta-casein precursor.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HANNOVERIAN; TISSUE=LACTATING MAMMARY GLAND;
RA Lenasi T., Rogelj I., Debeljak M., Dovc P.;
RT "Primary structure of equine beta-casein (b-CN) cDNA.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF214526; AAG43954.1;
DR InterPro: IPR001588; Casein.
DR Pfam: PF00363; caseins; 1.
DR PROSITE: PS00306; CASEIN_ALPHA_BETA; 1.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 1 15 POTENTIAL.
Qy SEQUENCE 233 AA; 26126 MW; 492A07BD010FBF88 CRC64;

Query Match
Best Local Similarity 49.5%; Score 52; DB 6; Length 233;
Matches 12; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 REELNVPGEIVEXLXXE 20
Db 16 REKEELNVSSTVESLSSE 35

RESULT 4
Q9MAR5
ID Q9MAR5 PRELIMINARY; PRT; 640 AA.
AC Q9MAR5;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

```

```

DE F28H19.8 protein.
GN F28H19.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eurosids II; Brassicales; Brassicaceae; Arabidopsi.
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altairi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler E., Chao Q., Chin C., Chiu J., Choi E., Gonzalez A.,
RA Hwang B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharsy N., Pham P., Sakano H., Shinn P.,
RA Toriumi M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC006423; AAF63114.1;
DR InterPro: IPR001584; Rve.
DR Pfam: PF00665; rve; 1.
DR PROSITE: PS00306; CASEIN_ALPHA_BETA; 1.
SQ SEQUENCE 640 AA; 73395 MW; 885DE941A980CD55 CRC64;

Query Match
Best Local Similarity 45.7%; Score 48; DB 10; Length 640;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 4 REELNVPGEIVEXLXXEERSIT 25
Db 70 REELNVPGEIVEXLXXEERSIT 91

RESULT 5
Q9XAW4
ID Q9XAW4 PRELIMINARY; PRT; 291 AA.
AC Q9XAW4;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DE Hypothetical 32.7 kDa protein.
OS Pseudomonas alcaligenes.
OG Plasmid pRA2.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX Pseudomonas.
OX NCBI_TaxID=43263;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NCIB 9867; TRANSPOSOM=TN5563;
RA MEDLINE=98126538; PubMed=94653390;
RA Kwong S.M., Yeo C.C., Chuah D., Poh C.L.;
RT "Sequence analysis of plasmid pRA2 from Pseudomonas alcaligenes NCIB
RT 9867 (P25X) reveals a novel replication region.";
RL FEMS Microbiol. Lett. 158:159-165(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=NCIB 9867; TRANSPOSOM=TN5563;
RA MEDLINE=98415121; PubMed=9742696;
RA Yeo C.C., Tham J.M., Kwong S.M., Ylin S., Poh C.L.;
RT "pTn5563, a transposon encoding putative mercuric ion transport
RT proteins located on plasmid pRA2 of Pseudomonas alcaligenes.";
RL FEMS Microbiol. Lett. 165:253-260(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=NCIB 9867; TRANSPOSOM=TN5563;
RA MEDLINE=20082846;
RA Kwong S.M., Yeo C.C., Suwanto A., Poh C.L.;
RT "Characterization of the endogenous plasmid from Pseudomonas
RT alcaligenes NCIB 9867: DNA sequence and mechanism of transfer.";
RL J. Bacteriol. 182:81-90(2000).
DR EMBL: U88088; AAD40349.1;
KW Hypothetical protein; plasmid.
SQ SEQUENCE 291 AA; 32661 MW; 970E473ADC8F479D CRC64;

Query Match
Best Local Similarity 44.8%; Score 47; DB 2; Length 291;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

```

```

Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 ELEELNVPGEIVEXLXXXESITR 25
    ||| : | : | : | : | : |
Db 161 ELESVRLSGDPLESLAGIDELIIR 184

RESULT 6
Q8YS03 ID Q8YS03 PRELIMINARY; PRT; 1289 AA.
AC Q8YS03;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein All3289.
GN ALL3289.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003592; BAB74988.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1289 AA; 141199 MW; FAE2808D4EFD7C4C CRC64;

Query Match 44.8%; Score 47; DB 16; Length 1289;
Best Local Similarity 69.2%; Pred. No. 69;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ELEELNVPGEIVE 14
    ||||| | : |
Db 609 ELEELNRPCTVLE 621

RESULT 7
Q9ZPS1 ID Q9ZPS1 PRELIMINARY; PRT; 334 AA.
AC Q9ZPS1;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE AT2902030 protein.
GN AT2902030.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Bonito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006532; AAD20095.1; -
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
DR PROSITE; PS0181; FBOX; 1.
SQ SEQUENCE 334 AA; 37978 MW; OD827279DD2EEF3F CRC64;

Query Match 43.8%; Score 46; DB 10; Length 334;
Best Local Similarity 37.0%; Pred. No. 23;
Matches 10; Conservative 8; Mismatches 7; Indels 2; Gaps 1;

QY 1 RELEELNVPGEIVEXLXXX--EESITR 25
    : | : : : | | | : | : | : |
Db 32 KEFEKIHIPNEIVEILVRLPKSLTR 58

RESULT 8
Q9C597 ID Q9C597 PRELIMINARY; PRT; 544 AA.
AC Q9C597;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE DNA excision repair protein.
GN AT5G21300.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL589883; CAC34483.1; -
DR InterPro; IPR01005; MYB_DNA_binding.
DR PROSITE; PS00037; MYB_1; UNKNOWN_1.
SQ SEQUENCE 544 AA; 61940 MW; 4B5E73E2B9EB75CF CRC64;

Query Match 43.3%; Score 45.5; DB 10; Length 544;
Best Local Similarity 38.5%; Pred. No. 47;
Matches 10; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

QY 1 RELEELNVPGEIVEXL-XXXESITR 25
    || : | | : | : | : |
Db 11 RENRRINTDGEAIDALFMATEAVTR 36

RESULT 9
O53142 ID O53142 PRELIMINARY; PRT; 488 AA.
AC O53142;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Thiol:fumarate reductase subunit B protein.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98237600; PubMed=9578488;
RA Helm S., Kunkel A., Thauer R.K., Hedderich R.;
RT "Thiol:fumarate reductase (tfr) from Methanobacterium

```

RT thermoautotrophicum Identification of the catalytic sites for fumarate
 reduction and thiol oxidation.",
 RL Eur. J. Biochem. 253:292-299(1998).
 CC -I- COFACTOR: BINDS A 2FE-2S CLUSTER (BY SIMILARITY).
 DR EMBL: AJ000942; CAA04399.1; -
 DR HSSP: P17596; 1OLB.
 DR InterPro: IPR000564; 2Fe2S_ferredoxin.
 DR InterPro: IPR001450; 4Fe4S_ferredoxin.
 DR InterPro: IPR004489; DhsB.
 DR InterPro: IPR004017; DUF224.
 DR InterPro: IPR001041; Ferredoxin.
 DR Pfam: PF02754; DUF224; 2.
 DR Pfam: PF00111; fer2; 1.
 DR TIGRfams: TIGR00384; dhsB; 1.
 DR PROSITE: PS00197; 2FE2S-FERREDOXIN; 1.
 DR PROSITE: PS00198; 4FE4S-FERREDOXIN; 2.
 KW Iron-sulfur.
 SQ SEQUENCE 488 AA; 53986 MW; 032EB355C4C8FEFA CRC64;

Query Match 42.9%; Score 45; DB 1; Length 488;
 Best Local Similarity 61.5%; Pred. No. 50;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 EELNVPGEIVEXL 16
 :|||||: :|
 Db 201 KELNVPGDALKEKL 213

RESULT 10

ID O27878 PRELIMINARY; PRT; 492 AA.
 AC O27878;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JUN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Fumarate reductase.
 GN MTH1850.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=DELTA H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
 Jiawani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 McDougall S., Shiner G., Goyal A., Petrovski S., Church G.M.,
 Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 "Complete genome sequence of Methanobacterium thermoautotrophicum
 deltaH: functional analysis and comparative genomics";
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -I- COFACTOR: BINDS A 2FE-2S CLUSTER (BY SIMILARITY).
 DR EMBL: AB000937; AAB86316.1; -
 DR HSSP: P17596; 1OLB.
 DR InterPro: IPR000564; 2Fe2S_ferredoxin.
 DR InterPro: IPR001450; 4Fe4S_ferredoxin.
 DR InterPro: IPR004489; DhsB.
 DR InterPro: IPR004017; DUF224.
 DR InterPro: IPR001041; Ferredoxin.
 DR Pfam: PF02754; DUF224; 2.
 DR Pfam: PF00111; fer2; 1.
 DR TIGRfams: TIGR00384; dhsB; 1.
 DR PROSITE: PS00197; 2FE2S-FERREDOXIN; 1.
 DR PROSITE: PS00198; 4FE4S-FERREDOXIN; 2.
 KW Iron-sulfur; Complete proteome.
 SQ SEQUENCE 492 AA; 54780 MW; 0EF9554ED909C5DB CRC64;

Query Match 42.9%; Score 45; DB 17; Length 492;
 Best Local Similarity 61.5%; Pred. No. 51;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 4 EELNVPGEIVEXL 16
 :|||||: :|
 Db 205 KELNVPGDALKEKL 217

RESULT 11

Q97WH8 PRELIMINARY; PRT; 587 AA.
 ID Q97WH8;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE BPS2 protein homolog (bps2).
 GN BPS2 OR SSO2241.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 Aveyez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
 De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
 Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.;
 "The complete genome of the crenarchaeon Sulfolobus solfataricus P2";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 RL EMBL: AE006828; AAK42408.1; -
 KW Complete proteome.
 SQ SEQUENCE 587 AA; 68436 MW; B90DBC1E19C05E86 CRC64;

Query Match 42.9%; Score 45; DB 17; Length 587;
 Best Local Similarity 45.5%; Pred. No. 62;
 Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
 QY 2 EELNVPGEIVEXLXXEESI 23
 ||:|:|:|:|:|:|:|
 Db 425 ELQLLGIPSSILEELKEKEHI 446

RESULT 12

Q9GZ12 PRELIMINARY; PRT; 807 AA.
 ID Q9GZ12;
 AC Q9GZ12;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 92.2 kDa protein.
 GN F52G3.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.;
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Blanchard M., Gattung S., Sansone J.;
 RT "The sequence of C. elegans cosmid F52G3.";
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

Wed Feb 12 11:35:22 2003

```
[3]
RN  SEQUENCE FROM N.A.
RP  Sampei G., Mizobuchi K.;
RC  STRAIN=BRISTOL N2;
RA  Waterston R.;
RT  "Direct Submission.";
RL  Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AF026212; AAF99971.1; -.
KW  Hypothetical protein.
SQ  SEQUENCE 807 AA; 92161 MW; 6919513906B0DB0D CRC64;

Query Match 42.9%; Score 45; DB 5; Length 807;
Best Local Similarity 45.0%; Pred. No. 88;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 ELEEINVPGEIVEXLXXEE 21
|||:|:|:|:|:|:|:|
Db 277 ELDELDSGNVPDGLTAE 296

RESULT 13
Q93GL4 PRELIMINARY; PRT; 1752 AA.
ID Q93GL4
AC Q93GL4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein.
DE Conjugative transfer: orit nicking-unwinding.
GN TRAI OR PSL108.
OS Salmonella typhimurium.
OC Bacteriophage.
OC Plasmid pSLT.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McLelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2."
RL Nature 413:852-856(2001).
DR EMBL: AE006471; AAL23509.1; -.
KW Plasmid; Complete proteome.
SQ SEQUENCE 1752 AA; 191683 MW; 6182AC451D5BE1F8 CRC64;

Query Match 42.9%; Score 45; DB 16; Length 1752;
Best Local Similarity 44.4%; Pred. No. 2,1e+02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 8 VPGEIVEXLXXEESITR 25
|||:|:|:|:|:|:|:|
Db 960 VPGELMETLISGQRAATR 977

RESULT 14
Q9WTB0 PRELIMINARY; PRT; 1756 AA.
ID Q9WTB0
AC Q9WTB0;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE DNA helicase I.
GN TRAI.
OS Escherichia coli.
OC Plasmid R100.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RT "Organization and diversification of plasmid genomes: complete
nucleotide sequence of the R100 genome.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF000342; BAA78885.1; -.
KW Helicase.
SQ SEQUENCE 1756 AA; 191683 MW; 56477ED2DC915BB3 CRC64;

Query Match 42.9%; Score 45; DB 2; Length 1756;
Best Local Similarity 44.4%; Pred. No. 2,1e+02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

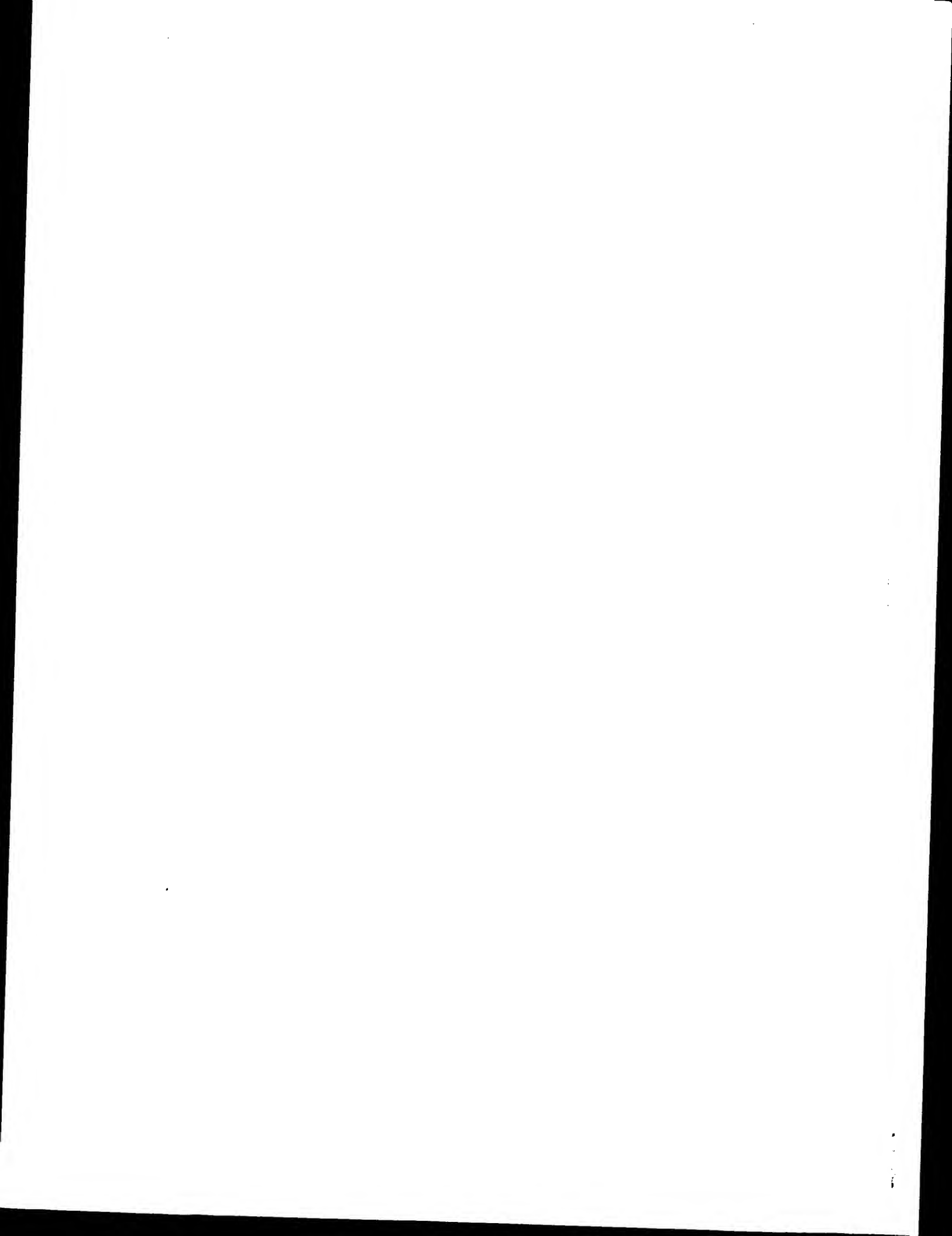
QY 8 VPGEIVEXLXXEESITR 25
|||:|:|:|:|:|:|:|
Db 961 VPGELMETLISGQRAATR 978

RESULT 15
Q9KZP7 PRELIMINARY; PRT; 284 AA.
ID Q9KZP7
AC Q9KZP7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein.
DE SC01385 OR SCL18A.05C.
GN SC01385 OR SCL18A.05C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Thomson N.R., Brown A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Harper D., Bateman A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Cronin A., Fraser A., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Huang C.H., Kieser T., Larke L., Rutherford K., Rutter S.,
RA Rabinowitz E., Rajandream M.A., Rotherford K., Taylor K.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL: AL353861; CAB88881.1; -.
KW Hypothetical protein.
SQ SEQUENCE 284 AA; 30476 MW; A29B3E2E9E021D97 CRC64;

Query Match 42.4%; Score 44.5; DB 16; Length 284;
Best Local Similarity 36.0%; Pred. No. 33;
Matches 9; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

QY 1 RELE-ELNVPGEIVEXLXXEESIT 24
::|:|:|:|:|:|:|:|
Db 234 QDLEPALNIPGGVQTLEKEQATV 258

Search completed: February 11, 2003, 18:19:42
Job time : 30.0928 secs
```



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:06:55 ; Search time 36.3402 Seconds
(without alignments)
91.669 Million cell updates/sec

Title: US-09-380-738A-3
Perfect score: 107
Sequence: 1 NANEEYSGXXXXAEVATEVK 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	96.3	25	8	Phosphopeptide 4.
2	103	96.3	25	12	Phosphopeptide #4
3	103	96.3	25	13	Anticariogenic pho
4	103	96.3	25	14	Casein phosphopept
5	103	96.3	25	14	Phosphopeptide der
6	103	96.3	25	15	Sequence of casein
7	103	96.3	25	16	Sodium caseinate t
8	103	96.3	25	19	Bos alpha-s2-casei
9	103	96.3	25	19	Phosphopeptide T4.
10	103	96.3	222	18	Bovine alpha-S2 ca

11	103	96.3	222	23	AAE17468	Bovine alpha-S2 ca
12	95	88.8	25	14	AAE17469	phosphopeptide 4.
13	88	82.2	223	23	AAE17469	Alpha-S2 casein pr
14	88	82.2	223	23	AAE17471	Capra hircus alpha
15	88	82.2	223	23	AAE17475	Alpha-S2 casein pr
16	84	78.5	223	23	AAE17470	Goat alpha-S2 case
17	84	78.5	223	23	AAE17473	Sheep alpha-S2 cas
18	50	46.7	1277	22	ABB71133	Drosophila melanog
19	46.5	43.5	1417	22	ABB55356	Drosophila melanog
20	46	43.0	783	22	ABB57036	Mouse ischaemic co
21	46	43.0	984	22	AAO12986	Human polyptide
22	44	41.1	189	21	AAG32328	Arabidopsis thalia
23	44	41.1	197	21	AAG32327	Arabidopsis thalia
24	44	41.1	197	23	ABB91821	Herbicidally activ
25	44	41.1	200	21	AAG05238	Arabidopsis thalia
26	44	41.1	200	21	AAG50050	Arabidopsis thalia
27	44	41.1	205	22	AA86168	Arabidopsis thalia
28	44	41.1	208	21	AA979386	Arabidopsis thalia
29	44	41.1	208	22	AA804153	Arabidopsis thalia
30	44	41.1	208	23	AAE21532	Arabidopsis thalia
31	44	41.1	208	23	AAE20526	Arabidopsis thalia
32	44	41.1	270	21	AAG05237	Arabidopsis thalia
33	44	41.1	270	21	AAG50049	Arabidopsis thalia
34	44	41.1	271	21	AAG05236	Arabidopsis thalia
35	44	41.1	271	21	AAG50048	Arabidopsis thalia
36	44	41.1	271	21	AA979392	Arabidopsis thalia
37	44	41.1	271	21	AA804159	Arabidopsis thalia
38	44	41.1	271	22	AA804159	Arabidopsis thalia
39	44	41.1	271	23	ABB79070	Arabidopsis thalia
40	44	41.1	271	23	AAE21538	Arabidopsis thalia
41	44	41.1	279	21	AAE20532	Arabidopsis thalia
42	44	41.1	279	21	AAE20532	Arabidopsis thalia
43	44	41.1	349	21	AAE20532	Arabidopsis thalia
44	44	41.1	350	21	AAE20532	Arabidopsis thalia
45	41	38.3	209	21	AAE20532	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAP71323
ID AAP71323 standard; peptide; 25 AA.
XX AAP71323;
XX AC
XX DT 07-MAY-1991 (first entry)
XX Phosphopeptide 4.
XX Caries; gingivitis; periodontal disease; osteoporosis; osteomalacia.
XX Synthetic.
XX Key Location/Qualifiers
XX Modified-site 11 /label= phosphoserine
XX Modified-site 12 /label= phosphoserine
XX Modified-site 13 /label= phosphoserine
XX Modified-site 16 /label= phosphoserine
XX W08707616-A.
XX 17-DEC-1987.
XX 12-JUN-1987; 87WO-AU00172.
XX 12-JUN-1986; 86AU-0006385.
XX (VICT-) VICTORIA DAIRY INDUSTRY AUTHORITY.
PA

QY 1 NANEESYSGXXEAEVATEVK 25
 |||||
 Db 1 NANEESYSGSSSEAEVATEVK 25

RESULT 6

AAR47817
 ID AAR47817 standard; peptide; 25 AA.
 XX
 AC AAR47817;
 XX
 DT 21-JUL-1994 (first entry)
 XX
 DE Sequence of casein phosphopeptide (CPP).
 XX
 KW Casein phosphopeptide; dental hypersensitivity; therapy; tooth.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 14
 FT /label= Ser(P)
 FT /note= "see also residues 16,17,18"
 XX
 PN W09400146-A.
 XX
 PD 06-JAN-1994.
 XX
 PF 29-JUN-1993; 93WO-AU00319.
 XX
 PR 29-JUN-1992; 92AU-0003221.
 XX
 PA (UYME) UNIV MELBOURNE.
 PA (VICT-) VICTORIAN DAIRY IND AUTHORITY.
 XX
 PI Reynolds EC;
 XX
 DR WPI; 1994-025888/03.
 XX

PT Treatment of dental hypersensitivity - using casein, component
 of casein, phospho-protein or phospho-peptide or their salts
 PS Disclosure; Page 14; 23pp; English.
 XX

CC Ser(P) = post-translationally phosphorylated serine. A mixture of
 casein phosphopeptides (CPP) and/or their salts may be used in a
 method for treating dental hypersensitivity. Pref. those CPPs
 contg. the sequence -Ser(P)-Ser(P)-predominate. The CPPs
 can be extracted from a casein digest.
 CC
 XX

SQ Sequence 25 AA;

Query Match 96.3%; Score 103; DB 15; Length 25;
 Best Local Similarity 84.0%; Pred. No. 7.1e-11;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NANEESYSGXXEAEVATEVK 25
 |||||
 Db 1 NANEESYSGSSSEAEVATEVK 25

RESULT 7

AAR68939
 ID AAR68939 standard; Peptide; 25 AA.
 XX
 AC AAR68939;
 XX

DT 07-SEP-1995 (first entry)

XX Sodium caseinate tryptic phosphopeptide T4.
 DE
 XX

KW Phosphopeptide; tryptic digestion; sodium caseinate; alpha(s1)-caseinate;

KW
 KW
 KW
 XX

phosphoserine; phosphothreonine; phosphotyrosine; phosphohistidine;
 sugar; sorbitol; mannitol; xylitol; lactitol; cellobiitol; confectionary;
 caries; gingivitis; calcium; remineralisation.

OS Bos taurus.

XX
 FH Key Location/Qualifiers
 FT Modified-site 11
 FT /label= OTHER
 FT /note= "Phosphoserine"
 FT
 FT Modified-site 12
 FT /label= OTHER
 FT /note= "Phosphoserine"
 FT
 FT Modified-site 13
 FT /label= OTHER
 FT /note= "Phosphoserine"
 FT
 FT Modified-site 16
 FT /label= OTHER
 FT /note= "Phosphoserine"

XX EP629393-A.

XX 21-DEC-1994.

XX 07-JUN-1994; 94EP-0304083.

XX 16-JUN-1993; 93US-0078706.

XX (ICIL) ICI AMERICAS INC.

XX Duross JW;

XX WPI; 1995-04845/07.

XX Crystalline poly:ol compositions for use in confectionery and
 pharmaceuticals - contain a phospho:peptide uniformly
 distributed within the crystal matrix and can improve oral
 hygiene

PS Disclosure; column 4; 8pp; English.

XX Peptides (AAR68936-40) are phosphopeptides resulting from the tryptic
 digestion of sodium caseinate. The peptide shown here is derived from
 the TCRK-tryptic digest of alpha(s2)-caseinate (comprising alpha(s2),
 alpha(s3), alpha(s4) and alpha(s6)). The peptides contain the amino
 acids phosphoserine, phosphothreonine, phosphotyrosine or
 phosphohistidine, and include in their sequence aspartate and glutamate.
 CC The phosphopeptides form part of a novel composition containing the
 peptide dispersed uniformly in a crystal mix of sugars chosen from
 sorbitol, mannitol, xylitol, lactitol, cellobiitol or mixtures of
 sorbitol/mannitol or sorbitol/xylitol. The compositions can be used
 instead of conventional sugars to manufacture confectionary products.
 CC The phosphopeptides are known to inhibit caries and gingivitis and can
 act as a source of calcium ions to promote/enhance remineralisation.

SQ Sequence 25 AA;

Query Match 96.3%; Score 103; DB 16; Length 25;
 Best Local Similarity 84.0%; Pred. No. 7.1e-11;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NANEESYSGXXEAEVATEVK 25
 |||||
 Db 1 NANEESYSGSSSEAEVATEVK 25

RESULT 8

AAW66601
 ID AAW66601 standard; peptide; 25 AA.
 XX
 AC AAW66601;
 XX

DT 27-NOV-1998 (first entry)

us-09-380-738a-3.rag

Wed Feb 12 11:35:22 2003

XX DE Bos alpha-s2-casein X-4P (f46-70) phosphopeptide.
XX AC casein; calcium phosphate complex; amorphous calcium phosphate; ACP;
XX DT phosphopeptide; delivery vehicle; calcium fluoride; calcium deficiency;
XX KW osteoporosis; osteomalacia; tooth; bone disease.
XX OS Synthetic.
XX OS Bos taurus.
XX FH Key Location/Qualifiers
XX FT Modified-site 11 /note= "Ser(P)"
XX FT Modified-site 12 /note= "Ser(P)"
XX FT Modified-site 13 /note= "Ser(P)"
XX FT Modified-site 16 /note= "Ser(P)"
XX FT W09840406-A1.
XX PN 17-SEP-1998.
XX PD 13-MAR-1998; 98WO-AU00160.
XX PF 13-MAR-1997; 97AU-0005662.
XX PR (UYME) UNIV MELBOURNE.
XX PA (VICT-) VICTORIAN DAIRY IND AUTHORITY.
XX PI Reynolds EC;
XX DR WPI; 1998-520803/44.
XX PT Stable calcium phosphate complex including phosphopeptide
XX PT stabilised amorphous calcium phosphate - useful for treatment of
XX PT dental caries, calcium malabsorption and bone diseases such as
XX PT osteoporosis and osteomalacia.
XX PS Claim 4; Page 35; 43pp; English.
XX CC The invention relates to a stable calcium phosphate complex including
XX CC phosphopeptide stabilised amorphous calcium phosphate (ACP) or its
XX CC derivative, where the phosphopeptide includes the amino acid sequence:
XX CC Ser(P)-Ser(P)-Glu-Glu. The amorphous phases stabilised by the
XX CC phosphopeptides are a delivery vehicle for co-localisation of Ca,
XX CC P and phosphate at the tooth surface in a slow-release amorphous form
XX CC producing superior anticaries efficacy over prior art. The amorphous
XX CC phases stabilised by the phosphopeptides are also useful as dietary
XX CC supplements to increase calcium bioavailability and to help prevent
XX CC diseases associated with calcium deficiencies. They are particularly
XX CC useful for treatment or prevention of dental caries, calcium
XX CC malabsorption and bone diseases such as osteoporosis and osteomalacia.
XX CC The compositions are useful in humans and in veterinary medicine in
XX CC domestic animals such as cattle, sheep, horses and companion animals e.g.
XX CC cats and dogs as well as zoo animals. The present sequence represents a
XX CC phosphopeptide component of a specifically claimed complex.
XX SQ Sequence 25 AA;
Query Match 96.3%; Score 103; DB 19; Length 25;
Best Local Similarity 84.0%; Pred. No. 7.1e-11;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 NANE EY SIG XXXX EAEVATEEVK 25
IIIIIIII II IIIIIII
DB 1 NANE EY SIG SSS EAEVATEEVK 25
RESULT 9
AAB12803
ID AAB12803 standard; peptide; 25 AA.

XX AC AAB12803;
XX DT 23-NOV-2000 (first entry)
XX DE Phosphopeptide T4.
XX KW Phosphopeptide; protective peptide formulation; thixotropic gel;
XX KW thickening agent; antimicrobial; antiinflammatory; osteopathic;
XX KW dental caries; dental sensitivity; gingivitis; mouth odour;
XX KW osteoporosis; calculus formation; oral cavity; recrystallising;
XX KW remineralising; enamel; dentine; buffering plaque.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 11 /note= "phosphoserine"
XX FT Modified-site 12 /note= "phosphoserine"
XX FT Modified-site 13 /note= "phosphoserine"
XX FT Modified-site 16 /note= "phosphoserine"
XX FT A0721548-B.
XX PN 06-JUL-2000.
XX PD 13-JAN-2000; 2000AU-0010070.
XX PF 13-JAN-1999; 99AU-0008135.
XX PR (PACI-) PACIFIC BIOLINK PTY LTD.
XX PA Bannister DJ;
XX PI WPI; 2000-482993/43.
XX DR Protective peptide formulation for treating and preventing various
XX DR dental disorders, gingivitis in humans and animals, has a suspension or
XX PT solution of phosphopeptide in water and thickening agent to form a gel
XX PT Claim 1; Page 18; 21pp; English.
XX CC The present invention describes a protective peptide formulation (I)
XX CC comprising a suspension or solution of a phosphopeptide (II) in water
XX CC and thickening agent to form a thixotropic gel. (I) can have
XX CC antiinflammatory and osteopathic activities. (I) is useful for treating
XX CC dental caries and dental sensitivity, gingivitis, mouth
XX CC odour, osteoporosis, calculus formation in oral cavity in humans and
XX CC animals. (I) is also useful for recrystallising and remineralising the
XX CC enamel and/or dentine and also for buffering plaque against decrease in
XX CC pH in humans and animals. Unlike prior art formulations, (I) has an
XX CC extended shelf life and is capable of producing a reliable gel without a
XX CC loss of activity of (II) and other bioactive constituents. AAB12800 to
XX CC AAB12804 represent specifically claimed phosphopeptides for use in the
XX CC present invention.
XX SQ Sequence 25 AA;
Query Match 96.3%; Score 103; DB 21; Length 25;
Best Local Similarity 84.0%; Pred. No. 7.1e-11;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 NANE EY SIG XXXX EAEVATEEVK 25
IIIIIIII II IIIIIII
DB 1 NANE EY SIG SSS EAEVATEEVK 25
RESULT 10
AAB12803
ID AAB12803 standard; peptide; 25 AA.

PA (UNIL) UNILEVER NV.
 PA (UNIL) UNILEVER PLC.
 XX Burger AR, Elliott DL, Schick LA;
 PI WPI; 1993-019802/03.
 XX Oral compens. contg. a phospho-peptide - with adn. of an anionic
 PT polymeric stabiliser to inhibit destabilisation in the oral
 PT environment
 XX Disclosure: Page 16; 18pp; English.
 XX The sequences given in AAR31237-42 represent phosphopeptides which
 CC were used as the active agents in an oral composition. These
 CC peptides were stabilised by an anionic polymeric stabiliser. The
 CC anionic polymers were chosen from a group consisting of carboxylate
 CC polymers, sulfonate polymers, polymers having both a carboxylate and
 CC a sulfonate moiety, and other such mixtures. The anionic polymeric
 CC stabiliser inhibits destabilisation of the phosphopeptide in the oral
 CC environment. These oral compositions, pref. containing a fluoride
 CC source may be used for inhibiting caries and gingivitis.
 XX Sequence 25 AA;
 SQ

Query Match 88.8%; Score 95; DB 14; Length 25;
 Best Local Similarity 80.0%; Pred. NO. 1.7e-09;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 NANESEYSGXXEAEVATEVK 25
 ||||| ||| || ||||| |||
 Db 1 NANESEYSGSSSEAEVATEVK 25

RESULT 13
 AAEL17469
 ID AAEL17469 standard; Protein: 223 AA.
 XX
 AC AAEL17469;
 XX
 DT 22-APR-2002 (first entry)
 XX
 DE Alpha-S2 casein precursor (alpha-S2-CN) protein #1.
 XX
 KW Alpha-S2 casein; alpha-S2-CN; dermatological; antiinflammatory; cosmetic;
 KW fibroblast; collagen; keratinocyte; skin regeneration; medicament; aging;
 KW toothpaste; chewing gum; cosmetic; wrinkling; periodontal disease.
 XX
 OS Unidentified.
 XX
 PN WO200202133-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 13-JUN-2001; 2001WO-GB02601.
 XX
 PR 30-JUN-2000; 2000GB-0016189.
 XX
 PA (PEPS-) PEPSYN LTD.
 XX
 PI Smith JA;
 XX
 XX WPI; 2002-154690/20.
 DR Use of peptide or its derivative containing an amino acid sequence in
 XX alpha-S2 casein precursor in the manufacture of a medicament for
 PT alleviating or preventing periodontal disease and an effect of aging in
 PT skin
 PT
 XX Disclosure: Page 8; 27pp; English.
 PS The invention relates to a composition comprising a peptide or its
 XX derivative. The peptide contains an amino acid sequence from alpha-S2
 CC casein precursor. The peptide stimulates the growth of fibroblasts, and
 CC thus the synthesis and secretion of collagen. The peptides also stimulate
 CC the growth of keratinocytes, which aid in the formation and regeneration
 CC of skin surface. The peptide is useful in the manufacture of a medicament
 CC in the form of a toothpaste or a chewing gum, for alleviating or
 CC preventing periodontal disease and a medicament in the form of a cosmetic
 CC composition for alleviating or preventing an effect of aging,
 CC particularly wrinkling of the skin. The present sequence is alpha-S2
 CC casein precursor (alpha-S2-CN) protein.
 XX
 SQ

CC casein precursor. The peptides stimulate the growth of fibroblasts, and
 CC thus the synthesis and secretion of collagen. The peptides also stimulate
 CC the growth of keratinocytes, which aid in the formation and regeneration
 CC of skin surface. The peptide is useful in the manufacture of a medicament
 CC in the form of a toothpaste or a chewing gum, for alleviating or
 CC preventing periodontal disease and a medicament in the form of a cosmetic
 CC composition for alleviating or preventing an effect of aging,
 CC particularly wrinkling of the skin. The present sequence is alpha-S2
 CC casein precursor (alpha-S2-CN) protein.
 XX
 SQ Sequence 223 AA;
 Query Match 82.2%; Score 88; DB 23; Length 223;
 Best Local Similarity 72.0%; Pred. NO. 3.2e-07;
 Matches 18; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 NANESEYSGXXEAEVATEVK 25
 ||||| ||| || ||||| |||
 Db 62 NANESEYSGSSSEAEVATEVK 86

RESULT 14
 AAEL17471
 ID AAEL17471 standard; Protein: 223 AA.
 XX
 AC AAEL17471;
 XX
 DT 22-APR-2002 (first entry)
 XX
 DE Capra hircus alpha-S2 casein C precursor protein.
 XX
 KW Goat; alpha-S2 casein; dermatological; antiinflammatory; fibroblast;
 KW collagen; keratinocyte; skin regeneration; medicament; toothpaste; aging;
 KW chewing gum; cosmetic; wrinkling; periodontal disease.
 XX
 OS Capra hircus.
 XX
 PN WO200202133-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 13-JUN-2001; 2001WO-GB02601.
 XX
 PR 30-JUN-2000; 2000GB-0016189.
 XX
 PA (PEPS-) PEPSYN LTD.
 XX
 PI Smith JA;
 XX
 XX WPI; 2002-154690/20.
 DR Use of peptide or its derivative containing an amino acid sequence in
 XX alpha-S2 casein precursor in the manufacture of a medicament for
 PT alleviating or preventing periodontal disease and an effect of aging in
 PT skin
 PT
 XX Claim 8; Page 8; 27pp; English.
 PS The invention relates to a composition comprising a peptide or its
 XX derivative. The peptide contains an amino acid sequence from alpha-S2
 CC casein precursor. The peptides stimulate the growth of fibroblasts, and
 CC thus the synthesis and secretion of collagen. The peptides also stimulate
 CC the growth of keratinocytes, which aid in the formation and regeneration
 CC of skin surface. The peptide is useful in the manufacture of a medicament
 CC in the form of a toothpaste or a chewing gum, for alleviating or
 CC preventing periodontal disease and a medicament in the form of a cosmetic
 CC composition for alleviating or preventing an effect of aging,
 CC particularly wrinkling of the skin. The present sequence is goat
 CC alpha-S2 casein C precursor protein.
 XX
 SQ Sequence 223 AA;
 Query Match 82.2%; Score 88; DB 23; Length 223;

Best Local Similarity 72.0%; Pred. No. 3.2e-07;
 Matches 18; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 NANESEYSGXXEAEVATEEVK 25
 |||||
 Db 62 NANESEYSGXXEAEVATEEVK 86

RESULT 15
 AAEL7475
 ID AAEL7475 standard; Protein; 223 AA.
 XX
 AC AAEL7475;
 XX
 DT 22-APR-2002 (first entry)
 XX
 DE Alpha-S2 casein precursor (alpha-S2-CN) protein #2.
 XX
 KW Alpha-S2 casein; alpha-S2-CN; dermatological; antiinflammatory; cosmetic;
 KW fibroblast; collagen; keratinocyte; skin regeneration; medication; aging;
 KW toothpaste; chewing gum; cosmetic; wrinkling; periodontal disease.
 XX
 OS Unidentified.
 XX
 PN WO200202133-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 13-JUN-2001; 2001WO-GB02601.
 XX
 PR 30-JUN-2000; 2000GB-0016189.
 XX
 PA (PEPS-) PEPSYN LTD.
 XX
 PI Smith JA;
 XX
 DR WPI; 2002-154690/20.
 XX
 PT Use of peptide or its derivative containing an amino acid sequence in
 PT alpha-S2 casein precursor in the manufacture of a medicament for
 PT alleviating or preventing periodontal disease and an effect of aging in
 PT skin -
 XX

PS Disclosure; Page 9; 27pp; English.
 XX
 CC The invention relates to a composition comprising a peptide or its
 CC derivative. The peptide contains an amino acid sequence from alpha-S2
 CC casein precursor. The peptides stimulate the growth of fibroblasts, and
 CC thus the synthesis and secretion of collagen. The peptides also stimulate
 CC the growth of keratinocytes, which aid in the formation and regeneration
 CC of skin surface. The peptide is useful in the manufacture of a medicament
 CC in the form of a toothpaste or a chewing gum, for alleviating or
 CC preventing periodontal disease and a medicament in the form of a cosmetic
 CC composition for alleviating or preventing an effect of aging, alpha-S2
 CC particularly wrinkling of the skin. The present sequence is alpha-S2
 CC casein precursor (alpha-S2-CN) protein.
 XX
 SQ Sequence 223 AA;

Query Match 82.2%; Score 88; DB 23; Length 223;
 Best Local Similarity 72.0%; Pred. No. 3.2e-07;
 Matches 18; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 NANESEYSGXXEAEVATEEVK 25
 |||||
 Db 62 NANESEYSGXXEAEVATEEVK 86

Search completed: February 11, 2003, 18:16:55
 Job time : 37.3402 secs

Wed Feb 12 11:35:22 2003

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:14:31 ; Search time 12,3711 seconds
(without alignments)
59.459 Million cell updates/sec

Title: US-09-380-738A-3

Perfect score: 107

Sequence: 1 NANEEYSGXXXXXAEVATEVK 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pap.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pap.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pap.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pap.*
- 5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pap.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	96.3	25	1	US-07-748-344B-4
2	103	96.3	25	2	US-08-954-985A-4
3	103	96.3	25	4	US-08-137-086-4
4	103	96.3	222	3	US-09-066-408-6
5	88	82.2	223	3	US-09-066-408-7
6	88	82.2	223	3	US-09-066-408-9
7	84	78.5	223	3	US-09-066-408-8
8	84	78.5	223	3	US-09-066-408-11
9	44	41.1	208	4	US-09-181-183-10
10	44	41.1	208	4	US-09-280-040-10
11	44	41.1	208	4	US-09-277-700-10
12	44	41.1	271	4	US-09-181-183-36
13	44	41.1	271	4	US-09-277-700-36
14	44	41.1	271	4	US-09-277-700-36
15	40	37.4	235	3	US-09-066-408-12
16	39.5	36.9	439	4	US-09-171-461-32
17	39	36.4	41	1	US-08-453-924-8
18	39	36.4	234	4	US-09-270-391-16
19	39	36.4	822	4	US-08-684-932A-38
20	38.5	36.0	246	3	US-09-238-481-2
21	38.5	36.0	246	4	US-09-572-810A-2
22	38	35.5	394	4	US-09-195-188-1
23	37.5	35.0	421	1	US-08-132-649-6
24	37.5	35.0	421	3	US-08-767-579-6
25	37.5	35.0	522	4	US-09-232-160-23
26	37	34.6	522	4	US-08-961-083-120
27	37	34.6	637	4	US-09-360-545-6

Sequence 58, Appl
Sequence 1, Appl
Sequence 31, Appl
Sequence 31, Appl
Sequence 118, Appl
Sequence 23, Appl
Sequence 24, Appl
Sequence 4947, Ap
Sequence 7, Appl
Sequence 2, Appl
Sequence 4825, Ap
Sequence 23, Appl
Sequence 40, Appl
Sequence 3, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-07-748-344B-4
; Sequence 4, Application US/07748344B
; Patent No. 5227154
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; TITLE OF INVENTION: PHOSPHOPEPTIDES FOR THE
; TREATMENT OF DENTAL CALCULUS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: ANDRUS, SCEALES, STARKE & SAWALL
; STREET: 100 EAST WISCONSIN AVE., SUITE 1100
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07748,344B
; FILING DATE: 19910822
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: SARA, CHARLES S
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: C.8493-87
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 255-2182
; TELEFAX: (608) 255-2182
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 11
; OTHER INFORMATION: Post-translationally phosphorylated serine
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: Phosphoserine

OTHER INFORMATION: Post-translationally phosphorylated serine
FEATURE: NAME/KEY: Phosphoserine
LOCATION: 16
OTHER INFORMATION: Post-translationally phosphorylated serine
US-08-137-086-4

Query Match 96.3%; Score 103; DB 4; Length 25;
Best Local Similarity 84.0%; Pred. No. 8e-12; 4; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 NANEEYSIGXXXEXAEVATEEVK 25
Db 1 NANEEYSIGSSSESAEVATEEVK 25

RESULT 4
US-09-066-408-6
; Sequence 6, Application US/09066408
; Patent No. 6060448
; GENERAL INFORMATION:
; APPLICANT: Smith, John Arthur
; APPLICANT: Wilkinson, Mark Charles
; APPLICANT: Liu, Qing-Ming
; TITLE OF INVENTION: Casein Fragments Having Growth Promoting
; TITLE OF INVENTION: Activity
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,408
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/02658
; FILING DATE: 31-OCT-1996
; APPLICATION NUMBER: GB 9522302.0
; FILING DATE: 31-OCT-1996
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..222
; OTHER INFORMATION: /note= "bovine alpha-s2 casein precursor"
US-09-066-408-6

Query Match 96.3%; Score 103; DB 3; Length 222;

Best Local Similarity 84.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 NANEEYSIGXXXEXAEVATEEVK 25
Db 61 NANEEYSIGSSSESAEVATEEVK 85

RESULT 5
US-09-066-408-7
; Sequence 7, Application US/09066408
; Patent No. 6060448
; GENERAL INFORMATION:
; APPLICANT: Smith, John Arthur
; APPLICANT: Wilkinson, Mark Charles
; APPLICANT: Liu, Qing-Ming
; TITLE OF INVENTION: Casein Fragments Having Growth Promoting
; TITLE OF INVENTION: Activity
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,408
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/02658
; FILING DATE: 31-OCT-1996
; APPLICATION NUMBER: GB 9522302.0
; FILING DATE: 31-OCT-1996
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..223
; OTHER INFORMATION: /note= "goat alpha-s2 casein precursor allele A"
US-09-066-408-7

Query Match 82.2%; Score 88; DB 3; Length 223;
Best Local Similarity 72.0%; Pred. No. 4.4e-08;
Matches 18; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 NANEEYSIGXXXEXAEVATEEVK 25
Db 62 NANEEYSIRSSSESAEVAPEIK 86

RESULT 6
US-09-066-408-9

```

; Sequence 9, Application US/09066408
; Patent No. 6060448
; GENERAL INFORMATION:
; APPLICANT: Smith, John Arthur
; APPLICANT: Wilkenson, Mark Charles
; APPLICANT: Liu, Qing-Ming
; TITLE OF INVENTION: Casein Fragments Having Growth Promoting
; TITLE OF INVENTION: Activity
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,408
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/02658
; FILING DATE: 31-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9522302.0
; FILING DATE: 31-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 018317-0001000S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..223
; OTHER INFORMATION: /note= "goat alpha-S2 casein
; OTHER INFORMATION: precursor allele C"
; US-09-066-408-9

Query Match 82.2%; Score 88; DB 3; Length 223;
Best Local Similarity 72.0%; Pred. No. 4.4e-08;
Matches 18; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 NANEESYICGXXEAEVATEVK 25
Db 62 NANEESYIRSSSESAKVAPEIK 86

RESULT 7
US-09-066-408-8
; Sequence 8, Application US/09066408
; Patent No. 6060448
; GENERAL INFORMATION:
; APPLICANT: Smith, John Arthur
; APPLICANT: Wilkenson, Mark Charles
; APPLICANT: Liu, Qing-Ming
; TITLE OF INVENTION: Casein Fragments Having Growth Promoting
; TITLE OF INVENTION: Activity
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,408
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/02658
; FILING DATE: 31-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9522302.0
; FILING DATE: 31-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 018317-0001000S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..223
; OTHER INFORMATION: /note= "goat alpha-S2 casein
; OTHER INFORMATION: precursor allele B"
; US-09-066-408-8

Query Match 78.5%; Score 84; DB 3; Length 223;
Best Local Similarity 68.0%; Pred. No. 2.3e-07;
Matches 17; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 NANEESYICGXXEAEVATEVK 25
Db 62 NANEESYIRSSSESAKVAPEIK 86

RESULT 8
US-09-066-408-11
; Sequence 11, Application US/09066408
; Patent No. 6060448
; GENERAL INFORMATION:
; APPLICANT: Smith, John Arthur
; APPLICANT: Wilkenson, Mark Charles
; APPLICANT: Liu, Qing-Ming
; TITLE OF INVENTION: Casein Fragments Having Growth Promoting
; TITLE OF INVENTION: Activity
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,408
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/02658
; FILING DATE: 31-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9522302.0
; FILING DATE: 31-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 018317-0001000S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..223
; OTHER INFORMATION: /note= "goat alpha-S2 casein
; OTHER INFORMATION: precursor allele B"
; US-09-066-408-8

```


Db 55 NAEFTVGLAPETLRKTSLEELK 77

RESULT 11

US-09-277-700-10
; Sequence 10, Application US/09277700
; Patent No. 6350597

; GENERAL INFORMATION:

; APPLICANT: VIITANEN, PAUL V.

; APPLICANT: BACOT, KAREN O.

; APPLICANT: JORDAN, DOUGLAS B.

; TITLE OF INVENTION: RIBOFLAVIN SYNTHASE GENES AND ENZYMES

; FILE REFERENCE: CL-1083-B

; CURRENT APPLICATION NUMBER: US/09/277,700

; CURRENT FILING DATE: 1999-03-26

; EARLIER APPLICATION NUMBER: 08/912,218

; EARLIER FILING DATE: AUGUST 15, 1997

; NUMBER OF SEQ ID NOS: 39

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 10

; LENGTH: 208

; TYPE: PRT

; ORGANISM: arabidopsis

US-09-277-700-10

Query Match 41.1%; Score 44; DB 4; Length 208;
Best Local Similarity 34.8%; Pred. No. 2.3;
Matches 8; Conservative 6; Mismatches 9; Indels 9; Gaps 0;

Qy 3 NEEYSIGXXXEXAEVATEEVK 25

Db 55 NAEFTVGLAPETLRKTSLEELK 77

RESULT 12

US-09-181-183-36

; Sequence 36, Application US/09181183

; Patent No. 6146866

; GENERAL INFORMATION:

; APPLICANT: VIITANEN, PAUL VEIKKO

; APPLICANT: BACOT, KAREN ONLEY

; APPLICANT: JORDAN, DOUGLAS BRIAN

; TITLE OF INVENTION: LUMAZINE SYNTHASE AND

; TITLE OF INVENTION: RIBOFLAVIN SYNTHASE

; NUMBER OF SEQUENCES: 39

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

; STREET: 1007 MARKET STREET

; CITY: WILMINGTON

; STATE: DELAWARE

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 19898

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.50 INCH

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95

; SOFTWARE: MICROSOFT WORD VERSION 7.0A

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/181,183

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: FLOYD, LINDA AXAMETHY

; REGISTRATION NUMBER: 33,692

; REFERENCE/DOCKET NUMBER: CL-1083

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 302-992-8112

; TELEFAX: 302-773-0164

; INFORMATION FOR SEQ ID NO: 36:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 271 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: arabidopsis RS precursor
US-09-181-183-36

Query Match 41.1%; Score 44; DB 4; Length 271;
Best Local Similarity 34.8%; Pred. No. 3.1;
Matches 8; Conservative 6; Mismatches 9; Indels 9; Gaps 0;

Qy 3 NEEYSIGXXXEXAEVATEEVK 25

Db 118 NAEFTVGLAPETLRKTSLEELK 140

RESULT 13

US-09-280-040-36

; Sequence 36, Application US/09280040

; Patent No. 6323013

; GENERAL INFORMATION:

; APPLICANT: VIITANEN, PAUL VEIKKO

; APPLICANT: BACOT, KAREN ONLEY

; APPLICANT: JORDAN, DOUGLAS BRIAN

; TITLE OF INVENTION: LUMAZINE SYNTHASE AND

; TITLE OF INVENTION: RIBOFLAVIN SYNTHASE

; NUMBER OF SEQUENCES: 39

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

; STREET: 1007 MARKET STREET

; CITY: WILMINGTON

; STATE: DELAWARE

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 19898

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.50 INCH

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95

; SOFTWARE: MICROSOFT WORD VERSION 7.0A

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/280,040

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: FLOYD, LINDA AXAMETHY

; REGISTRATION NUMBER: 33,692

; REFERENCE/DOCKET NUMBER: CL-1083

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 302-992-8112

; TELEFAX: 302-773-0164

; INFORMATION FOR SEQ ID NO: 36:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 271 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; INDIVIDUAL ISOLATE: arabidopsis RS precursor

US-09-280-040-36

Query Match 41.1%; Score 44; DB 4; Length 271;

Best Local Similarity 34.8%; Pred. No. 3.1;

Matches 8; Conservative 6; Mismatches 9; Indels 9; Gaps 0;

RESULT 14

US-09-277-700-36

; Sequence 36, Application US/09277700

Best Local Similarity	34.8%	Pred: No. 3.1;		
		6. Mismatches	9: Indels	0: Gaps

```

01 3 NEEEYSGXXXXXAEVATEVK 25
02 | | | | | | : : | | | |
03 118 NAEFTVGLAPETLRKYSLEELK 140
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
10
```


;
; INDIVIDUAL ISOLATE: Mature arabidopsis RS
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-870-756-10

Query Match 41.1%; Score 44; DB 10; Length 208;
Best Local Similarity 34.8%; Pred. No. 1.9;
Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 3 NEEYSIGXXXEXAEVATEEVK 25
| | | | | : | : | : |
Db 55 NAEFTVGLAPETLRKTSLEELK 77

RESULT 2

US-09-874-585B-10
; Sequence 10, Application US/09874585B
; Patent No. US20020127670A1

GENERAL INFORMATION:

; APPLICANT: Viitanen, Paul Veikko
; APPLICANT: Jordan, Karen Onley
; TITLE OF INVENTION: Riboflavin Synthase Genes and Enzymes and Methods of Use
; FILE REFERENCE: CL1083 US DIV2
; CURRENT APPLICATION NUMBER: US/09/874,585B
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/874,585
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: 08/912,218
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
; LENGTH: 208
; TYPE: PRT
; ORGANISM: arabidopsis

US-09-874-585B-10

Query Match 41.1%; Score 44; DB 10; Length 208;
Best Local Similarity 34.8%; Pred. No. 1.9;
Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 3 NEEYSIGXXXEXAEVATEEVK 25
| | | | | : | : | : |
Db 55 NAEFTVGLAPETLRKTSLEELK 77

RESULT 3

US-09-870-756-36
; Sequence 36, Application US/09870756
; Patent No. US20020052023A1

GENERAL INFORMATION:

; APPLICANT: VIITANEN, PAUL VEIKKO
; APPLICANT: JORDAN, KAREN ONLEY
; TITLE OF INVENTION: LUMAZINE SYNTHASE AND
; RIBOFLAVIN SYNTHASE
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898

COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/870,756
; FILING DATE: 31-May-2001
; CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AYAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1083
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: arabidopsis RS precursor
; SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-870-756-36

Query Match 41.1%; Score 44; DB 10; Length 271;
Best Local Similarity 34.8%; Pred. No. 2.5;
Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 3 NEEYSIGXXXEXAEVATEEVK 25
| | | | | : | : | : |
Db 118 NAEFTVGLAPETLRKTSLEELK 140

RESULT 4

US-09-874-585B-36
; Sequence 36, Application US/09874585B
; Patent No. US20020127670A1

GENERAL INFORMATION:

; APPLICANT: Viitanen, Paul Veikko
; APPLICANT: Jordan, Karen Onley
; APPLICANT: Bacot, Douglas Brain
; TITLE OF INVENTION: Riboflavin Synthase Genes and Enzymes and Methods of Use
; FILE REFERENCE: CL1083 US DIV2
; CURRENT APPLICATION NUMBER: US/09/874,585B
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/874,585
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: 08/912,218
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 36
; LENGTH: 271
; TYPE: PRT
; ORGANISM: arabidopsis
US-09-874-585B-36

Query Match 41.1%; Score 44; DB 10; Length 271;
Best Local Similarity 34.8%; Pred. No. 2.5;
Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 3 NEEYSIGXXXEXAEVATEEVK 25
| | | | | : | : | : |
Db 118 NAEFTVGLAPETLRKTSLEELK 140

RESULT 5

US-09-970-711-32
; Sequence 32, Application US/09970711
; Patent No. US20020081279A1

GENERAL INFORMATION:

; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chioccia, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800001

1. TITLE OF INVENTION: Identification of Essential Genes in

ORGANISM: *Staphylococcus aureus*

; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-12030

Query Match 35.5%; Score 38; DB 10; Length 103;
Best Local Similarity 43.5%; Pred. No. 9.2;
Matches 10; Conservative 2; Mismatches 11; Indels 0; Gaps 0;
QY 2 ANEEYSIGXXXEAATEEV 24
Db 42 ANGEDVKIGLPVVEGAKVTAEVV 64

Search completed: February 11, 2003, 18:36:15
Job time : 13.1134 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:17:01 ; Search time 171.649 Seconds
(without alignments)
93.903 Million cell updates/sec

Title: US-09-380-738A-3

Perfect score: 107

Sequence: 1 NANEEYSIGXXEAEVATEEVK 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/PTCUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	103	96.3	25	3	US-07-731-592B-4
2	99	92.5	25	17	US-09-380-738A-3
3	50	46.7	1277	20	US-09-614-150-40191
4	50	46.7	1277	27	US-60-191-681-39826
5	50	46.7	1277	27	US-60-191-681-30853
6	48	44.9	137	21	US-09-791-537-140550

7	47	43.9	1280	21	US-09-791-537-137326
8	46.5	43.5	581	27	US-60-173-464-18812
9	46.5	43.5	764	27	US-60-161-932-2609
10	46.5	43.5	1417	20	US-09-614-150-22860
11	46.5	43.5	1417	27	US-60-191-637-22940
12	46.5	43.5	1417	27	US-60-191-681-18089
13	46.5	43.5	1454	21	US-09-791-537-12245
14	46	43.0	139	18	US-09-497-191-756
15	46	43.0	179	20	US-09-675-784A-10079
16	46	43.0	375	21	US-09-758-449-1162
17	46	43.0	375	21	US-09-760-443-1469
18	46	43.0	375	26	US-10-212-054-1469
19	46	43.0	375	26	US-10-212-778-1162
20	46	43.0	526	21	US-09-791-537-141293
21	46	43.0	528	21	US-09-791-537-142582
22	46	43.0	562	21	US-09-791-537-141454
23	46	43.0	565	21	US-09-791-537-141473
24	46	43.0	565	21	US-09-791-537-141475
25	46	43.0	696	21	US-09-791-537-142094
26	46	43.0	737	21	US-09-791-537-142243
27	46	43.0	738	21	US-09-791-537-142246
28	46	43.0	748	21	US-09-791-537-142267
29	46	43.0	755	21	US-09-791-537-142287
30	46	43.0	766	21	US-09-791-537-142307
31	46	43.0	767	21	US-09-791-537-142309
32	46	43.0	768	21	US-09-791-537-142311
33	46	43.0	768	21	US-09-791-537-142312
34	46	43.0	768	21	US-09-791-537-142313
35	46	43.0	775	21	US-09-791-537-142338
36	46	43.0	776	21	US-09-791-537-142340
37	46	43.0	777	21	US-09-791-537-142342
38	46	43.0	777	21	US-09-791-537-142343
39	46	43.0	777	21	US-09-791-537-142344
40	46	43.0	777	21	US-09-791-537-142345
41	46	43.0	777	21	US-09-791-537-142346
42	46	43.0	777	21	US-09-791-537-145969
43	46	43.0	778	21	US-09-791-537-142349
44	46	43.0	778	21	US-09-791-537-142350
45	46	43.0	779	21	US-09-791-537-142355

ALIGNMENTS

RESULT 1

US-07-731-592B-4
Sequence 4, Application US/07731592B
GENERAL INFORMATION:
APPLICANT: Burger, Allan R.
APPLICANT: Elliott, David L.
APPLICANT: Schick, Laura A.
TITLE OF INVENTION: Oral Compositions Containing a Phosphopeptide
TITLE OF INVENTION: Phosphopeptide
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: Patent Department, Unilever United States, Inc.
STREET: 45 River Road
CITY: Edgewater
STATE: New Jersey
COUNTRY: USA
ZIP: 07020
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-3.5 inch, 1.44 Mb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word(R)
CURRENT APPLICATION NUMBER: US/07731,592B
FILING DATE: 19910717
CLASSIFICATION: 514
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 25

Sequence 137326, A
Sequence 18812, A
Sequence 2609, A
Sequence 22860, A
Sequence 22940, A
Sequence 18089, A
Sequence 12245, A
Sequence 756, App
Sequence 10079, A
Sequence 1162, A
Sequence 1469, A
Sequence 1469, A
Sequence 1162, A
Sequence 141293, A
Sequence 142582, A
Sequence 141454, A
Sequence 141473, A
Sequence 141475, A
Sequence 142094, A
Sequence 142243, A
Sequence 142246, A
Sequence 142267, A
Sequence 142287, A
Sequence 142307, A
Sequence 142309, A
Sequence 142311, A
Sequence 142312, A
Sequence 142313, A
Sequence 142338, A
Sequence 142340, A
Sequence 142342, A
Sequence 142343, A
Sequence 142344, A
Sequence 142345, A
Sequence 142346, A
Sequence 145969, A
Sequence 142349, A
Sequence 142350, A
Sequence 142355, A

```
; TYPE: AMINO ACID
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 11
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 12
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 13
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 16
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
US-07-731-592B-4
```

Query Match 96.3%; Score 103; DB 3; Length 25;
Best Local Similarity 84.0%; Pred. No. 1.1e-10;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 NANEESYSGXXEAEVATEVK 25
||||||| |||||||
Db 1 NANEESYSGSSSEAEVATEVK 25
```

RESULT 2

```
US-09-380-738A-3
; Sequence 3, Application US/09380738A
; GENERAL INFORMATION:
```

```
; APPLICANT: REYNOLDS, Eric
; TITLE OF INVENTION: CALCIUM PHOSPHOPEPTIDE COMPLEXES
; FILE REFERENCE: 040268/0161
; CURRENT APPLICATION NUMBER: US/09/380, 738A
; CURRENT FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: PCT/AU98/00160
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: AU P05662
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Bos sp.
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: Xaa is a phosphorylated Serine
; NAME/KEY: misc_feature
; LOCATION: (12)..(12)
; OTHER INFORMATION: Xaa is a phosphorylated Serine
; NAME/KEY: misc_feature
; LOCATION: (13)..(13)
; OTHER INFORMATION: Xaa is a phosphorylated Serine
; NAME/KEY: misc_feature
; LOCATION: (16)..(16)
; OTHER INFORMATION: Xaa is a phosphorylated Serine
US-09-380-738A-3
```

Query Match 92.5%; Score 99; DB 17; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.7e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 NANEESYSGXXEAEVATEVK 25
```

```
Db 1 NANEESYSGXXEAEVATEVK 25
||||||| ||||||| ||||||| |||||||
```

RESULT 3

```
US-09-614-150-40191
; Sequence 40191, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150
; CURRENT FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40191
; LENGTH: 1277
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-09-614-150-40191
```

Query Match 46.7%; Score 50; DB 20; Length 1277;
Best Local Similarity 47.6%; Pred. No. 48;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```
QY 2 ANEEYSIGXXEAEVATE 22
|:::| |::| |::|
Db 207 ADEDDYDGDDEVAEPSTE 227
```

RESULT 4

```
US-60-191-637-39826
; Sequence 39826, Application US/60191637
; GENERAL INFORMATION:
```

```
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39826
; LENGTH: 1277
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-637-39826
```

Query Match 46.7%; Score 50; DB 27; Length 1277;
Best Local Similarity 47.6%; Pred. No. 48;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```
QY 2 ANEEYSIGXXEAEVATE 22
```

-Wed Feb 12 11:35:23 2003

```

; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 137326
; LENGTH: 1280
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-791-537-137326

Query Match      43.9%; Score 47; DB 21; Length 1280;
Best Local Similarity 40.9%; Pred. No. 1.6e+02;
Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY      3 NEEYSIGXXXXXAEVATEE 24
      ::::||||| |::| |::|
Db      36 DDEYSIGTETSESKVEREKV 57
      ::::||||| |::| |::|

RESULT 8
US-60-173-464-18812
; Sequence 18812, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; FILE REFERENCE: CL000173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18812
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Drosophila
US-60-173-464-18812

Query Match      43.5%; Score 46.5; DB 27; Length 581;
Best Local Similarity 50.0%; Pred. No. 70;
Matches 13; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY      1 NANEET--YSIGXXXXXAEVATEE 23
      ||||| |::| |::| |::|
Db      225 NAEETEGSAAIEDAEETTEATEE 250
      ||||| |::| |::| |::|

RESULT 9
US-60-161-932-2609
; Sequence 2609, Application US/60161932
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic
; TITLE OF INVENTION: Acid Sequences, Systems Containin the Nucleic Acid Sequences
; FILE REFERENCE: CL000122
; CURRENT APPLICATION NUMBER: US/60/161,932
; CURRENT FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 2626
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2609
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Drosophila
US-60-161-932-2609

Query Match      43.5%; Score 46.5; DB 27; Length 764;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 13; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY      1 NANEET--YSIGXXXXXAEVATEE 23
      ||||| |::| |::| |::|
Db      342 NAEETEGSAAIEDAEETTEATEE 367
      ||||| |::| |::| |::|

; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 137326
; LENGTH: 1280
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-791-537-137326

Query Match      44.9%; Score 48; DB 21; Length 137;
Best Local Similarity 45.0%; Pred. No. 5.5;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      4 EEYSIGXXXXXAEVATEE 23
      ||||| |::| |::| |::|
Db      104 EEFNIGMAEKAQKIATVE 123
      ||||| |::| |::| |::|

RESULT 7
US-09-791-537-137326
; Sequence 137326, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 140550
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-791-537-140550

Query Match      46.7%; Score 50; DB 27; Length 1277;
Best Local Similarity 47.6%; Pred. No. 48;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2 ANEEYSIGXXXXXAEVATE 22
      |:::| |::| |::| |::|
Db      207 ADEDDYDGDDEEVAEPSTE 227
      |:::| |::| |::| |::|

RESULT 6
US-09-791-537-140550
; Sequence 140550, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 140550
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-791-537-140550

Query Match      46.7%; Score 50; DB 27; Length 1277;
Best Local Similarity 47.6%; Pred. No. 48;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2 ANEEYSIGXXXXXAEVATE 22
      |:::| |::| |::| |::|
Db      207 ADEDDYDGDDEEVAEPSTE 227
      |:::| |::| |::| |::|

RESULT 5
US-60-191-681-30853
; Sequence 30853, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: c1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30853
; LENGTH: 1277
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-681-30853

Query Match      46.7%; Score 50; DB 27; Length 1277;
Best Local Similarity 47.6%; Pred. No. 48;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2 ANEEYSIGXXXXXAEVATE 22
      |:::| |::| |::| |::|
Db      207 ADEDDYDGDDEEVAEPSTE 227
      |:::| |::| |::| |::|

RESULT 4
US-60-191-681-30853
; Sequence 30853, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: c1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30853
; LENGTH: 1277
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-681-30853

Query Match      46.7%; Score 50; DB 27; Length 1277;
Best Local Similarity 47.6%; Pred. No. 48;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2 ANEEYSIGXXXXXAEVATE 22
      |:::| |::| |::| |::|
Db      207 ADEDDYDGDDEEVAEPSTE 227
      |:::| |::| |::| |::|

RESULT 3
US-60-191-681-30853
; Sequence 30853, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: c1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30853
; LENGTH: 1277
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-681-30853

Query Match      46.7%; Score 50; DB 27; Length 1277;
Best Local Similarity 47.6%; Pred. No. 48;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2 ANEEYSIGXXXXXAEVATE 22
      |:::| |::| |::| |::|
Db      207 ADEDDYDGDDEEVAEPSTE 227
      |:::| |::| |::| |::|

RESULT 2
US-60-191-681-30853
; Sequence 30853, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: c1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30853
; LENGTH: 1277
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-681-30853

Query Match      46.7%; Score 50; DB 27; Length 1277;
Best Local Similarity 47.6%; Pred. No. 48;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2 ANEEYSIGXXXXXAEVATE 22
      |:::| |::| |::| |::|
Db      207 ADEDDYDGDDEEVAEPSTE 227
      |:::| |::| |::| |::|

RESULT 1
US-60-191-681-30853
; Sequence 30853, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: c1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30853
; LENGTH: 1277
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-681-30853

Query Match      46.7%; Score 50; DB 27; Length 1277;
Best Local Similarity 47.6%; Pred. No. 48;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY      2 ANEEYSIGXXXXXAEVATE 22
      |:::| |::| |::| |::|
Db      207 ADEDDYDGDDEEVAEPSTE 227
      |:::| |::| |::| |::|

```

```

RESULT 10
US-09-614-150-22860
; Sequence 22860, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150
; CURRENT FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22860
; LENGTH: 1417
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-09-614-150-22860

Query Match          43.5%; Score 46.5; DB 20; Length 1417;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 13; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 1 NANEET---YSIGXXXEXAEVATEE 23
   ||||| :|| |||||
Db 225 NAEETEGSAAIEDAEETTEATEE 250

RESULT 11
US-60-191-637-22940
; Sequence 22940, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22940
; LENGTH: 1417
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-637-22940

Query Match          43.5%; Score 46.5; DB 27; Length 1417;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 13; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 1 NANEET---YSIGXXXEXAEVATEE 23
   ||||| :|| |||||
Db 225 NAEETEGSAAIEDAEETTEATEE 250

RESULT 12
US-60-191-681-18089
; Sequence 18089, Application US/60191681
; GENERAL INFORMATION:
; APPLICANT: Li, Peter, W.D.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING ISOLATED DROSOPHILA PROTEINS AND
; TITLE OF INVENTION: USES THEREOF.
; FILE REFERENCE: C1000390
; CURRENT APPLICATION NUMBER: US/60/191,681
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 30973
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18089
; LENGTH: 1417
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-191-681-18089

Query Match          43.5%; Score 46.5; DB 27; Length 1417;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 13; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 1 NANEET---YSIGXXXEXAEVATEE 23
   ||||| :|| |||||
Db 225 NAEETEGSAAIEDAEETTEATEE 250

RESULT 13
US-09-791-537-12245
; Sequence 12245, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12245
; LENGTH: 1454
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-791-537-12245

Query Match          43.5%; Score 46.5; DB 21; Length 1454;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 13; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 1 NANEET---YSIGXXXEXAEVATEE 23
   ||||| :|| |||||
Db 225 NAEETEGSAAIEDAEETTEATEE 250

RESULT 14
US-09-497-191-756
; Sequence 756, Application US/09497191
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: CHEN, Xianfeng
; APPLICANT: SUBRAMANIAN, Gopalakrishnan
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS 84-89, 91-99
; FILE REFERENCE: 2750-694P
; CURRENT APPLICATION NUMBER: US/09/497,191
; CURRENT FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 920
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 756
; LENGTH: 139

```

us-09-380-738a-3.rapm

--Wed Feb 12 11:35:23 2003

TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: any n or xaa = unknown
US-09-497-191-756

Query Match 43.0%; Score 46; DB 18; Length 139;
Best Local Similarity 36.0%; Pred. No. 13;
Matches 9; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 1 NANESEYSGXXEAEVATEVK 25
|||||:|:|:|:|:|:|:
Db 79 NANESEFKIGTSGVKQKAENNIK 103

RESULT 15

US-09-675-784A-10079
; Sequence 10079, Application US/09675784A
; GENERAL INFORMATION:
; APPLICANT: HARE, ROBERTA S.
; APPLICANT: SHAW, KAREN J.
; APPLICANT: SHIMER JR., GEORGE H.
; APPLICANT: KESSLER, MARCO
; APPLICANT: NOLLING, JORK
; APPLICANT: ZENG, QIANDONG
; APPLICANT: GREENE, JONATHAN R.
; TITLE OF INVENTION: ASPERGILLUS FUMIGATUS NUCLEIC ACIDS AND POLYPEPTIDES,
; FILE REFERENCE: 2976-40200SI
; CURRENT APPLICATION NUMBER: US/09/675,784A
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/156,338
; PRIOR FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 13925
; SEQ ID NO 10079
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-675-784A-10079

Query Match 43.0%; Score 46; DB 20; Length 179;
Best Local Similarity 34.8%; Pred. No. 18;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 NANESEYSGXXEAEVATEE 23
|||||:|:|:|:|:|:|:
Db 140 NANESEYSGXXEAEVATEE 162

Search completed: February 11, 2003, 18:33:23
Job time : 173.649 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:17:56 ; Search time 28.6082 Seconds
(without alignments)
72.734 Million cell updates/sec

Title: US-09-380-738A-3

Perfect score: 107

Sequence: 1 NANEEYSIGXXXXEAEVATEEVK 25

Scoring table: BLOSUM62

Gap: 10.0 , Gapext 0.5

Searched: 420401 seqs, 83231269 residues

Total number of hits satisfying chosen parameters: 420401

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match, 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New: *
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pap: *
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pap: *
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pap: *
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pap: *
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pap: *
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pap: *
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	43.0	779	6	US-10-222-851-28
2	43	40.2	1286	6	US-10-310-154-618
3	40	37.4	137	1	PCT-US02-25586-24
4	40	37.4	137	6	US-10-217-939-24
5	40	37.4	369	1	PCT-US02-36122-62
6	39	36.4	134	5	US-09-782-130-5
7	39	36.4	510	6	US-10-070-634-5
8	39	36.4	643	5	US-09-950-084-6839
9	39	36.4	830	6	US-10-325-899-9358
10	39	36.4	905	6	US-10-295-981-52
11	39	36.4	908	1	PCT-US02-37618-9
12	39	36.4	908	6	US-10-306-686-15
13	38.5	36.0	166	5	US-09-950-084-4560
14	38	35.5	212	5	US-09-724-676-92228
15	38	35.5	212	5	US-09-724-676-92233
16	38	35.5	212	5	US-09-724-676A-92228
17	38	35.5	212	5	US-09-724-676A-92233
18	38	35.5	458	5	US-09-636-458-8
19	38	35.5	461	5	US-09-724-676-76998
20	38	35.5	461	5	US-09-724-676A-76998
21	38	35.5	483	5	US-09-724-676-76999
22	38	35.5	483	5	US-09-724-676A-76999
23	38	35.5	488	1	PCT-US02-41115-11
24	38	35.5	492	5	US-09-636-458-2
25	38	35.5	492	5	US-09-636-458-4
26	38	35.5	492	5	US-09-636-458-6

27 38 35.5 492 5 US-09-636-458-10 Sequence 10, Appl
28 38 35.5 510 5 US-09-724-676-92227 Sequence 92227, A
29 38 35.5 510 5 US-09-724-676-92232 Sequence 92232, A
30 38 35.5 510 5 US-09-724-676A-92227 Sequence 92227, A
31 38 35.5 510 5 US-09-724-676A-92232 Sequence 92232, A
32 38 35.5 587 5 US-09-724-676-76994 Sequence 76994, A
33 38 35.5 587 5 US-09-724-676A-77004 Sequence 77004, A
34 38 35.5 587 5 US-09-724-676A-76994 Sequence 76994, A
35 38 35.5 587 5 US-09-724-676A-77004 Sequence 77004, A
36 38 35.5 609 5 US-09-724-676-76995 Sequence 76995, A
37 38 35.5 609 5 US-09-724-676-77005 Sequence 77005, A
38 38 35.5 609 5 US-09-724-676A-76995 Sequence 76995, A
39 38 35.5 609 5 US-09-724-676A-77005 Sequence 77005, A
40 38 35.5 664 5 US-09-724-676-77009 Sequence 77009, A
41 38 35.5 664 5 US-09-724-676A-77009 Sequence 77009, A
42 38 35.5 678 5 US-09-724-676-77008 Sequence 77008, A
43 38 35.5 678 5 US-09-724-676A-77008 Sequence 77008, A
44 38 35.5 811 6 US-10-310-154-512 Sequence 512, App
45 38 35.5 1385 5 US-09-724-676-53883 Sequence 53883, A

ALIGNMENTS

RESULT 1
US-10-222-851-28
; Sequence 28, Application US/10222851
; GENERAL INFORMATION:
; APPLICANT: Hitachi LTD.
; TITLE OF INVENTION: RP42-INTERACTING PROTEINS AND USE THEREOF
; FILE REFERENCE: H0100926A1
; CURRENT APPLICATION NUMBER: US/10/222,851
; CURRENT FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-851-28

Query Match 43.0%; Score 46; DB 6; Length 779;
Best Local Similarity 47.6%; Pred. No. 9;
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 EEEYSIGXXXXEAEVATEEV 24
||| | | | : |||
Db 316 EEEETGNSSEAEQSAEEV 336

RESULT 2
US-10-310-154-618
; Sequence 618, Application US/10310154
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S
; APPLICANT: Adams, Thomas H
; APPLICANT: Ruif, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Bodupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jinhua
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihshieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent

; APPLICANT: Kretzmer, Keith A.
 ; APPLICANT: Laccetti, Lucille B.
 ; APPLICANT: Lai, Chao-Qiang
 ; APPLICANT: Lee, Gary
 ; APPLICANT: Lin, Jie-Yi
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Lu, Bin
 ; APPLICANT: Luethy, Michael M.
 ; APPLICANT: Lund, Adrian
 ; APPLICANT: Madson, Linda L.
 ; APPLICANT: Malloy, Kathleen A.
 ; APPLICANT: McKiel, Christine L.
 ; APPLICANT: Miller, Philip W.
 ; APPLICANT: Padmavathi, Manchikanti
 ; APPLICANT: Parnell, Laurence D.
 ; APPLICANT: Start, William G.
 ; APPLICANT: Vidya, K.R.
 ; APPLICANT: Wang, Haiyun
 ; APPLICANT: Xin, Zhanqiao
 ; APPLICANT: Xu, Nanfei
 ; APPLICANT: Yang, Chunzhi
 ; APPLICANT: Zeng, Xiaoping
 ; APPLICANT: Zhang, Qiang
 ; APPLICANT: Zhao, Yajuan
 ; APPLICANT: Zhou, Li
 ; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
 ; FILE REFERENCE: 38-15(52796)B
 ; CURRENT APPLICATION NUMBER: US/10/310,154
 ; CURRENT FILING DATE: 2002-12-04
 ; PRIOR APPLICATION NUMBER: 60/337,358
 ; PRIOR FILING DATE: 2001-12-04
 ; NUMBER OF SEQ ID NOS: 736
 ; SEQ ID NO 618
 ; LENGTH: 1286
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 US-10-310-154-618

Query Match 40.2%; Score 43; DB 6; Length 1286;
 Best Local Similarity 55.0%; Pred. No. 52;
 Matches 11; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 4 EEEYSIGXXXXXAEVATEE 23
 ||||| ||| |||||
 Db 246 EEEKSAKSISEIMEATEE 265

RESULT 3
 PCT-US02-25586-24
 ; Sequence 24, Application PC/TUS0225586
 ; GENERAL INFORMATION:
 ; APPLICANT: BASF PLANT SCIENCE GMBH
 ; TITLE OF INVENTION: SUGAR AND LIPID METABOLISM REGULATORS IN PLANTS III
 ; FILE REFERENCE: 16313-0156
 ; CURRENT APPLICATION NUMBER: PCT/US02/25586
 ; CURRENT FILING DATE: 2002-08-12
 ; PRIOR APPLICATION NUMBER: 60/311,414
 ; PRIOR FILING DATE: 2001-08-10
 ; NUMBER OF SEQ ID NOS: 67
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 24
 ; LENGTH: 137
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 PCT-US02-25586-24

Query Match 37.4%; Score 40; DB 1; Length 137;
 Best Local Similarity 40.0%; Pred. No. 12;
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 QY 4 EEEYSIGXXXXXAEVATEE 23
 ||||| ||| |||||

Db 104 EEEFNQMAEEKAQKIATVE 123
 RESULT 4
 US-10-217-939-24
 ; Sequence 24, Application US/10217939
 ; GENERAL INFORMATION:
 ; APPLICANT: MITTENDORF, VOLKER
 ; APPLICANT: HAERTEL, HEIKO A.
 ; APPLICANT: CIRPUS, PETRA
 ; TITLE OF INVENTION: SUGAR AND LIPID METABOLISM REGULATORS IN PLANTS III
 ; FILE REFERENCE: 16313-0157
 ; CURRENT APPLICATION NUMBER: US/10/217,939
 ; CURRENT FILING DATE: 2002-08-12
 ; PRIOR APPLICATION NUMBER: 60/311,414
 ; PRIOR FILING DATE: 2001-08-10
 ; NUMBER OF SEQ ID NOS: 67
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 24
 ; LENGTH: 137
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-10-217-939-24

Query Match 37.4%; Score 40; DB 6; Length 137;
 Best Local Similarity 40.0%; Pred. No. 12;
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 QY 4 EEEYSIGXXXXXAEVATEE 23
 ||||| ||| |||||

Db 104 EEEFNQMAEEKAQKIATVE 123

RESULT 5
 PCT-US02-36122-62
 ; Sequence 62, Application PC/TUS0236122
 ; GENERAL INFORMATION:
 ; APPLICANT: Murphy, Ellen and Projan, Stephen, J.
 ; TITLE OF INVENTION: Alloococcus Otitidis Infectious Disease Targets
 ; FILE REFERENCE: Application 1
 ; CURRENT APPLICATION NUMBER: PCT/US02/36122
 ; CURRENT FILING DATE: 2002-11-25
 ; NUMBER OF SEQ ID NOS: 106
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 62
 ; LENGTH: 369
 ; TYPE: PRT
 ; ORGANISM: Alloococcus otitidis
 PCT-US02-36122-62

Query Match 37.4%; Score 40; DB 1; Length 369;
 Best Local Similarity 38.1%; Pred. No. 39;
 Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
 QY 4 EEEYSIGXXXXXAEVATEEV 24
 ||||| ||| ||| |||

Db 116 EEEVALAKRIEEDGDIKQEL 136

RESULT 6
 US-09-782-130-5
 ; Sequence 5, Application US/09782130
 ; GENERAL INFORMATION:
 ; APPLICANT: KNAUF, VIC C.
 ; APPLICANT: KRIDL, JEAN C.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR REGULATED TRANSCRIPTION
 ; FILE REFERENCE: 16518.052
 ; CURRENT APPLICATION NUMBER: US/09/782,130
 ; CURRENT FILING DATE: 2001-02-12
 ; PRIOR APPLICATION NUMBER: US 09/232,861
 ; PRIOR FILING DATE: 1999-01-15
 ; PRIOR APPLICATION NUMBER: US 08/812,665


```

> RESULT 3
> US-10-325-899-9358
> Sequence 9358, Application US/10325899
> GENERAL INFORMATION:
> APPLICANT: Wohlgemuth, Jay
> APPLICANT: Fry, Kirk
> APPLICANT: Ly, Ngoc
> APPLICANT: Woodward, Robert
> TITLE OF INVENTION: METHODS AND COMPO
> TITLE OF INVENTION: REJECTION
> FILE REFERENCE: 506612000122
> CURRENT APPLICATION NUMBER: US/10/325
> CURRENT FILING DATE: 2002-12-20
> PRIOR APPLICATION NUMBER: US 60/296,7
> PRIOR FILING DATE: 2001-06-08
> PRIOR APPLICATION NUMBER: US 10/006,2
> PRIOR FILING DATE: 2001-10-22
> PRIOR APPLICATION NUMBER: US 10/131,8
> PRIOR FILING DATE: 2002-04-24
> NUMBER OF SEQ ID NOS: 9966
> SOFTWARE: PatentIn version 3.1
> SEQ ID NO 9358
> LENGTH: 830

```

Query Match 36.4%; Score 39; DB 6; Length 830;
Best Local Similarity 39.1%;
Matches 9; Conservative 5; Mismatches 9; Indels

QY 2 ANEEYSIGXXXEEAEVATEEV 24
::||:| |||||
Db 690 SSDEKAGGLDDEEEAELVPSEV 712

US-10-295-981-52
; Sequence 52, Application US/10295981
; GENERAL INFORMATION;

INVENTOR: Bertin, John
 TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
 FILE REFERENCE: 07334-124001
 CURRENT APPLICATION NUMBER: US/10/295,981
 CURRENT FILING DATE: 2002-11-15

PRIOR APPLICATION NUMBER: US 09/340,620
 PRIOR FILING DATE: 1999-06-28
 PRIOR APPLICATION NUMBER: US 09/245,281
 PRIOR FILING DATE: 1999-02-05
 PRIOR APPLICATION NUMBER: US 09/207,359
 PRIOR FILING DATE: 1998-12-08
 PRIOR APPLICATION NUMBER: US 09/099,041
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: US 09/019,942
 PRIOR FILING DATE: 1998-02-06
 NUMBER OF SEQ ID NOS: 71

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 905
; TYPE: PRN
; ORGANISM: Rattus rattus
US-10-295-981-52

```

Query Match	36.4%;	Score 39;	DB 6;	Length 905;
Best Local Similarity	35.0%;	Pred. No. 1.6e+02;		
Matches	7;	Conservative	6;	Mismatches 7;
				Indels

QY 4 EEYSIGXXXEEEXAEVATEE 23
| : : | | : : | :
Db 112 EDPFSLGTITPEIAELSEEK 131

RESULT 11
PCT-US02-37618-9
; Sequence 9, Application PC/TUS0237618
: GENERAL INFORMATION

GENERAL INFORMATION:
APPLICANT: CANFIELD, William
TITLE OF INVENTION: METHODS OF PRODUCING HIGH MANNOSE GLYCOPROTEINS IN COMPLEX CARBOHYDRATE
TITLE OF INVENTION: DEFICIENT CELLS
FILE REFERENCE: 2035120577
CURRENT APPLICATION NUMBER: PCT/US02/37618
CURRENT FILING DATE: 2002-12-19
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.1

TYPE: PRT
 ORGANISM: Mus musculus
 CT-US02-37618-9

Query Match	36.4%	Score 39;	DB 1;	Length 908;
Best Local Similarity	43.5%	Pred. No. 1.7e+02;		
Matches 10; Conservative		3. Mismatched		

QY 1 NANEYSIGXXXEAEVATEE 23
|||:|:| : ||| |

Db 620 NANDEEFKI ---QIAVEVDTRE 638

US-10-306-686-15 ; Sequence 15, Application US/10306686 ; GENERAL INFORMATION:

```

/ APPLICANT: CAMFIELD, WILLIAM
/ TITLE OF INVENTION: PHOSDOSTER ALPHA-G
/ FILE REFERENCE: 230397US77DIV
/ CURRENT APPLICATION NUMBER: US/10/306.686
/ CURRENT FILING DATE: 2002-11-29
/ PRIOR APPLICATION NUMBER: 09/636,596
/ PRIOR FILING DATE: 2000-08-10
/ PRIOR APPLICATION NUMBER: 60/153,831
/ PRIOR FILING DATE: 1999-08-14
/ NUMBER OF SEQ ID NOS: 52
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 15
/ LENGTH: 908
/ TYPE: PRT
/ ORGANISM: Mus musculus
/ US-10-306-686-15

```

Query Match	36.48;	Score 39;	DB 6;	Length 908;
Best Local Similarity	43.58;	Pred. No. 1.7e+02;		
Matches 10:	Conservative	3. Mismatch		

```

QY      1 NANEESIGXXXEXAEVATEE 23
      III:III:I : III I
Db      620 NANDEEFKI---QIAVEVDTRE 630

```

US-09-950-084-4560

Sequence 4564, Application US/09950084
GENERAL INFORMATION:
APPLICANT: George H. Shimer, Jr.
APPLICANT: George H. Miller
APPLICANT: Roberta S. Hare
APPLICANT: Karen J. Shaw

TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
 FILE REFERENCE: 103471C963052
 CURRENT APPLICATION NUMBER: US/09/950,084
 CURRENT FILING DATE: 2001-09-10
 PRIOR APPLICATION NUMBER: US 09/417,811
 PRIOR FILING DATE: 1999-10-14
 PRIOR APPLICATION NUMBER: US 09/353,718

```

1 PRIOR FILING DATE: 1999-07-14
2 PRIOR APPLICATION NUMBER: US 09/266,557
3
4 PRIOR FILING DATE: 1999-03-11
5 PRIOR APPLICATION NUMBER: US 09/266,556
6
7 PRIOR FILING DATE: 1999-03-11
8 PRIOR APPLICATION NUMBER: US 09/266,555
9
10 PRIOR FILING DATE: 1999-03-11
11 PRIOR APPLICATION NUMBER: US 09/266,542
12
13 PRIOR FILING DATE: 1999-03-11
14 PRIOR APPLICATION NUMBER: US 09/266,541
15
16 PRIOR FILING DATE: 1999-03-11
17 PRIOR APPLICATION NUMBER: US 09/037,934
18
19 PRIOR FILING DATE: 1998-03-10
20 PRIOR APPLICATION NUMBER: US 09/036,720
21
22 PRIOR FILING DATE: 1998-03-06
23 PRIOR APPLICATION NUMBER: US 09/036,338
24
25 PRIOR FILING DATE: 1998-03-06
26 Remaining Prior Application data removed
27 NUMBER OF SEQ ID NOS: 7451

```

```

; SEQ ID NO 4560
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-950-084-4560

```

Wed Feb 12 11:35:23 2003 us-09-380-738a-3.rapn

Query Match 36.0%; Score 38.5; DB 5; Length 166;
 Best Local Similarity 36.0%; Pred. No. 28;
 Matches 9; Conservative 4; Mismatches 3; Gaps 1;

OY 4 EEEYSIG--XXEEXAEVATEEVK 25
 ||||| :||| |||
 DB 29 EGYNAVNYAGSKEKEAEVVEIK 53

RESULT 14
 US-09-724-676-92228
 ; Sequence 92228, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 92228
 ; LENGTH: 212
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-724-676-92228

Query Match 35.5%; Score 38; DB 5; Length 212;
 Best Local Similarity 35.5%; Pred. No. 45;
 Matches 11; Conservative 2; Mismatches 8; Indels 10; Gaps 1;

OY 5 EEEYSIGXXEEXAE-----VATEEVK 25
 ||||| ||| |||
 DB 134 EDYFSGAGEEEEEELQRLPMLSLTEDVK 164

RESULT 15
 US-09-724-676-92233
 ; Sequence 92233, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 92233
 ; LENGTH: 212
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-724-676-92233

Query Match 35.5%; Score 38; DB 5; Length 212;
 Best Local Similarity 35.5%; Pred. No. 45;
 Matches 11; Conservative 2; Mismatches 8; Indels 10; Gaps 1;

OY 5 EEEYSIGXXEEXAE-----VATEEVK 25
 ||||| ||| |||
 DB 134 EDYFSGAGEEEEEELQRLPMLSLTEDVK 164

Search completed: February 11, 2003, 18:35:19
 Job time : 28.6082 secs

Wed Feb 12 11:35:23 2003

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:14:11 : Search time 21.9072 Seconds
(without alignments)
109.706 Million cell updates/sec

Title: US-09-380-738A-3
Perfect score: 107
Sequence: 1 NANEETISIGXXEAEVATEEVK 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	103	96.3	222	1 KABOS2	alpha-s2-casein pr
2	88	82.2	223	2 I46995	alpha s2-casein C
3	84	78.5	223	1 KASHS2	alpha-s2-casein pr
4	84	78.5	223	2 JN0547	hypothetical prote
5	47	43.9	1280	2 G96796	hypothetical prote
6	46.5	43.5	1454	2 T13709	diacylglycerol kin
7	46	43.0	768	2 H54024	protein kinase (EC
8	46	43.0	777	2 B54024	protein kinase (EC
9	46	43.0	777	2 F54024	protein kinase (EC
10	46	43.0	779	2 E54024	protein kinase (EC
11	46	43.0	783	2 A55817	cyclin-dependent k
12	45	42.1	772	2 I50463	protein kinase - c
13	44	41.1	197	2 B84592	probable riboflavi
14	44	41.1	671	2 S53407	CHS5 protein - yea
15	43	40.2	88	2 S09659	hypothetical prote
16	43	40.2	250	2 T26010	hypothetical prote
17	43	40.2	332	2 I40726	rep protein - Cory
18	43	40.2	639	2 D70198	cell division prot
19	43	40.2	1287	2 S55954	viral mRNA transla
20	42	39.3	82	1 QFBO	micro glutamic aci
21	42	39.3	527	2 T16022	hypothetical prote
22	42	39.3	554	2 JW0094	neurofilament prot
23	42	39.3	562	2 A47710	metalloproteinase
24	42	39.3	881	2 S67026	probable late embr
25	41.5	38.8	136	2 T14305	conserved hypothet
26	41.5	38.8	516	2 T40181	proteasome beta-1
27	41	38.3	243	2 T48677	hypothetical prote
28	41	38.3	244	2 E84510	hypothetical prote
29	41	38.3	273	2 A11593	hypothetical prote

30	41	38.3	332	2 B72421	oligopeptide ABC t
31	41	38.3	471	2 T49019	probable RNA bindi
32	41	38.3	487	2 S60675	hypothetical prote
33	41	38.3	548	1 QFPG	neurofilament trip
34	41	38.3	561	2 F75191	hypothetical prote
35	41	38.3	562	2 I40227	neutral proteinase
36	41	38.3	696	2 S44912	hypothetical prote
37	41	38.3	1869	2 A59290	filamentous hemagg
38	41	38.3	4152	2 T31102	acyl carrier prote
39	40	37.4	137	2 S03267	alpha s2-casein -
40	40	37.4	235	2 A48383	NBP2 protein - yea
41	40	37.4	236	2 S57986	probable dnaJ-like
42	40	37.4	256	2 A81370	hypothetical prote
43	40	37.4	320	2 T21040	hypothetical prote
44	40	37.4	328	2 T01225	GRR1 related prote
45	40	37.4	364	2 T48752	

ALIGNMENTS

RESULT 1

KABOS2

alpha-s2-casein precursor - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 01-Sep-1981 #sequence_revision 03-Feb-1994 #text_change 22-Jun-1999

C:Accession: JQ2008; A29087; A91438; S66626; A03107

R:Groenen, M.A.M.; Dijkhof, R.J.M.; Verstege, A.J.M.; van der Poel, J.J.

Gene 123, 187-193, 1993

A:Title: The complete sequence of the gene encoding bovine alphas2-casein.

A:Reference number: JQ2008; MUID:93154583; PMID:8428658

A:Accession: JQ2008

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-222 <GRO>

A:Cross-references: GB:M94327

R:Stewart, A.F.; Bonsing, J.; Beattie, C.W.; Shah, F.; Willis, I.M.; Mackinlay, A.G.

Mol. Biol. Evol. 4, 231-241, 1987

A:Title: Complete nucleotide sequences of bovine alpha-s2- and beta-casein cDNAs: 15

A:Reference number: A93062; MUID:8818989; PMID:2833669

A:Accession: A29087

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-222 <STP>

A:Cross-references: GB:M16644; NID:q162928; PIDN:AAA30479.1; PID:q162929

R:Brignon, G.; Ribadeau Dumas, B.; Mercier, J.C.; Pelissier, J.P.; Das, B.C.

FEBS Lett. 76, 274-279, 1977

A:Title: Complete amino acid sequence of bovine alpha-S2-casein.

A:Reference number: A91438; MUID:77185633; PMID:862906

A:Contents: A allele

A:Accession: A91438

A:Molecule type: protein

A:Residues: 16-101, EE, 104-222 <BRI>

A>Note: four fractions, previously designated s2, s3, s4, and s6, appear to have t

these

R:Grosclaude, F.; Joudrier, P.; Mahe, M.F.

J. Dairy Res. 46, 211-213, 1979

A:Title: A genetic and biochemical analysis of a polymorphism of bovine alpha-s2-c

A:Reference number: A92777; MUID:79239837; PMID:469044

A:Contents: annotation; D allele

A>Note: the sequence of the D allele has a deletion of nine residues, which may be

R:Zucht, H.D.; Raids, M.; Adermann, K.; Maegert, H.J.; Forssmann, W.G.

FEBS Lett. 372, 185-188, 1995

A:Title: Casocidin-I: a casein-alpha(s2) derived peptide exhibits antibacterial ac

A:Reference number: S66626; MUID:9600204; PMID:7556666

A:Accession: S66626

A:Molecule type: protein

A:Residues: 165-203 <ZUC>

C:Comment: The sequence of the A allele is shown.

C:Genetics:

A:Gene: alphas2ca

A:Map position: 6

A:Introns: 17/3; 26/3; 33/3; 47/3; 56/3; 65/3; 74/3; 82/3; 97/3; 138/3; 147/3; 151/3

2

C:Genetics:
A:Gene: F28016.18
A:Map position: 1

Query Match 43.9%; Score 47; DB 2; Length 1280;
Best Local Similarity 40.9%; Pred. No. 11;
Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 3 NEEYSIGXXXXXAEVATEEV 24
:::||||| |::| |:
Db 36 DDDEYSIGTSELSESKVEEKV 57

RESULT 6

TI3709
diacylglycerol kinase (EC 2.7.1.107) - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: T13709
R:Masai, I.; Okazaki, A.; Hosoya, T.; Hotta, Y.
Proc. Natl. Acad. Sci. U.S.A. 90, 11157-11161, 1993
A:Title: Drosophila retinal degeneration A gene encodes an eye-specific diacylglycerol kinase
A:Reference number: Z17702; MUID:94068563; PMID:8248222
A:Accession: T13709
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1454 <MAS>
A:Cross-references: EMBL:DI7315; NID:di1007485; PID:di1004654; PIDN:BAA04135.1
C:Genetics:
A:Gene: rdgA
A:Cross-references: FlyBase:FBgn0003217
C:Keywords: phosphotransferase

Query Match 43.5%; Score 46.5; DB 2; Length 1454;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 13; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 1 NANEE---YSIGXXXXXAEVATEE 23
||||| |::| |:
Db 225 NAEDEEGSAAIEDAEETTEATEE 250

RESULT 7

H54024
protein kinase (EC 2.7.1.37) cdc2-related PITSLRE alpha 2-3 - human
C:Species: Homo sapiens (man)
C:Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 17-Nov-2000
C:Accession: H54024; G54024
R:Xiang, J.; Lahti, J.M.; Grenet, J.; Easton, J.; Kidd, V.J.
J. Biol. Chem. 269, 15786-15794, 1994
A:Title: Molecular cloning and expression of alternatively spliced PITSLRE protein kinases
A:Reference number: A54024; MUID:94253170; PMID:8195233
A:Accession: H54024
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-768 <XIA>
A:Cross-references: GB:U04817
A:Note: authors translated the codon AAA for residue 5 as Leu, GAC for residue 6 as Lys, CTC for residue 342 as Phe, CTG for residue 574 as Gln, AAG for residue 614 as Asn, GAT
A:Accession: G54024
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-114,117-768 <XI2>
A:Cross-references: GB:U07705
A:Note: authors translated the codon AAG for residue 5 as Leu, GAC for residue 6 as Lys, CTC for residue 351 as Phe, GAT for residue 418 as Glu, AAC for residue 457 as Lys, GTA
C:Keywords: alternative splicing; ATP; phosphotransferase
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
F:409-664/Domain: protein kinase homology <KIN>
F:417-425/Region: protein kinase ATP-binding motif

Query Match 43.0%; Score 46; DB 2; Length 768;
Best Local Similarity 47.6%; Pred. No. 9.6;

Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 BEEYSIGXXXXXAEVATEEV 24
||| |::| |:
Db 305 BEEETGNSERASQSAEEV 325

RESULT 8

B54024
protein kinase (EC 2.7.1.37) cdc2-related PITSLRE alpha 2-2 - human
N:Contains: protein kinase cdc2-related PITSLRE alpha 2-4
C:Species: Homo sapiens (man)
C:Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 19-Dec-1997
C:Accession: B54024; C54024
R:Xiang, J.; Lahti, J.M.; Grenet, J.; Easton, J.; Kidd, V.J.
J. Biol. Chem. 269, 15786-15794, 1994
A:Title: Molecular cloning and expression of alternatively spliced PITSLRE protein kinases
A:Reference number: A54024; MUID:94253170; PMID:8195233
A:Accession: B54024
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-777 <XIA>
A:Cross-references: GB:U04816
A:Accession: C54024
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 216-777 <XI2>
A:Cross-references: GB:U04818
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
C:Keywords: alternative splicing; ATP; phosphotransferase
F:418-673/Domain: protein kinase homology <KIN>
F:426-434/Region: protein kinase ATP-binding motif

Query Match 43.0%; Score 46; DB 2; Length 777;
Best Local Similarity 47.6%; Pred. No. 9.8;
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 BEEYSIGXXXXXAEVATEEV 24
||| |::| |:
Db 314 BEEETGNSERASQSAEEV 334

RESULT 9

F54024
protein kinase (EC 2.7.1.37) cdc2-related PITSLRE beta 2-1 - human
N:Contains: protein kinase cdc2-related PITSLRE beta-1
C:Species: Homo sapiens (man)
C:Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 19-Dec-1997
C:Accession: F54024; D54024
R:Xiang, J.; Lahti, J.M.; Grenet, J.; Easton, J.; Kidd, V.J.
J. Biol. Chem. 269, 15786-15794, 1994
A:Title: Molecular cloning and expression of alternatively spliced PITSLRE protein kinases
A:Reference number: A54024; MUID:94253170; PMID:8195233
A:Accession: F54024
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-777 <XIA>
A:Cross-references: GB:U07704
A:Accession: D54024
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 339-777 <XI2>
A:Cross-references: GB:U04819
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
C:Keywords: alternative splicing; ATP; phosphotransferase
F:418-673/Domain: protein kinase homology <KIN>
F:426-434/Region: protein kinase ATP-binding motif

Query Match 43.0%; Score 46; DB 2; Length 777;
Best Local Similarity 47.6%; Pred. No. 9.8;
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 BEEYSIGXXXXXAEVATEEV 24

us-09-380-738a-3.rpr

Wed Feb 12 11:35:23 2003

Query Match 41.1%; Score 44; DB 2; Length 671;
 Best Local Similarity 55.0%; Pred. No. 18;
 Matches 11; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3 NEEYSIGXXXEEAEVATE 22
 ||||| | |||||
 Db 278 NEEELSYKNEPVAEVADE 297

RESULT 15

S09659
 hypothetical protein A3 - Rhizobium sp. insertion sequence ISR1
 C:Species: Rhizobium sp.
 C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 08-Oct-1999
 C:Accession: S09659
 R:Primer: U.B.: Kalinowski, J.; Rueger, B.; Heumann, W.; Puehler, A.
 Plasmid 21, 120-128, 1989
 A:Title: ISR1, a transposable DNA sequence resident in Rhizobium class IV strains, shows
 A:Reference number: S09659; MUID:8929722; PMID:2544911
 A:Accession: S09659
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-88 <PRI>
 A:Cross-references: EMBL:X06616; NID:948670; PIDN:CAA29830.1; PID:948671
 C:Genetics:
 A:Mobile element: insertion sequence ISR1

Query Match 40.2%; Score 43; DB 2; Length 88;
 Best Local Similarity 55.0%; Pred. No. 2.9;
 Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 EYSIGXXXEEAEVATEEV 24
 || || | | || ||
 Db 8 EEQIIGILREQEAGVATAEV 27

Search completed: February 11, 2003, 18:21:16
 Job time : 23.9072 secs

Wed Feb 12 11:35:23 2003

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:07:35 ; Search time 11.0825 seconds
(without alignments)
93.563 Million cell updates/sec

Title: US-09-380-738a-3
Perfect score: 107
Sequence: 1 NANEYSIGXXXEAQVATEVK 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	96.3	222	1 CAS2_BOVIN	P02663 bos taurus
2	88	82.2	223	1 CAS2_CAPI	P33049 capra hircu
3	84	78.5	223	1 CAS2_SHEEP	P04654 ovis aries
4	46.5	43.5	1454	1 KDGE_DROME	Q09103 drosophila
5	44	41.1	671	1 CHS5_YEAST	Q12114 saccharomyc
6	43	40.2	88	1 YIA3_RHISP	P17985 rhizobium s
7	43	40.2	1287	1 SK12_YEAST	P35207 saccharomyc
8	42	39.3	554	1 NFL_BOVIN	P02548 bos taurus
9	42	39.3	562	1 NPMR_BACME	Q00891 bacillus me
10	42	39.3	881	1 ARP8_YEAST	Q12386 saccharomyc
11	41.5	38.8	516	1 YB4F_SCHPO	Q14360 schizosacch
12	41	38.3	548	1 NFL_PIG	P02547 sus scrofa
13	41	38.3	696	1 Y012_CABEL	P34668 caenorhabdi
14	40.5	37.9	193	1 CAS2_CAMDR	Q09794 camelus dro
15	40	37.4	137	1 ACP1_ARATH	P11829 arabidopsis
16	40	37.4	235	1 CAS2_PIG	P39036 sus scrofa
17	40	37.4	401	1 ENO_THEAC	Q01042 herpesvirus
18	40	37.4	407	1 IB6H_HSVSA	Q91171 thermoplasma
19	40	37.4	505	1 SPKD_SYNF3	P54735 synecocyst
20	40	37.4	660	1 VBI_HPV29	P50764 human papil
21	40	37.4	929	1 YDM6_SCHPO	P87137 schizosacch
22	40	37.4	988	1 E412_MOUSE	Q90318 mus musculu
23	40	37.4	1828	1 MY5A_RAT	Q90413 rattus norv
24	40	37.4	1853	1 MY5A_MOUSE	Q99104 mus musculu
25	39.5	36.9	439	1 PIV2_ADEGL	Q64749 avian adeno
26	39.5	36.9	520	1 RECN_AQUAE	O66834 aquifex aeo
27	39	36.4	134	1 ACP2_BRANA	P17650 brassica na
28	39	36.4	134	1 ACP3_BRANA	P32887 brassica na
29	39	36.4	134	1 ACP5_BRANA	P08971 brassica na
30	39	36.4	134	1 ACP_BRACM	P07088 brassica ca
31	39	36.4	415	1 PIC1_HUMAN	Q9nrds homo sapien
32	39	36.4	625	1 GIDA_STEAM	Q99qt4 staphylococ
33	39	36.4	830	1 VPP3_HUMAN	Q13488 h vacuolar

34	39	36.4	1386	1 ZAP3_MOUSE	Q9r017 mus musculu
35	38	35.5	136	1 ACP2_ARATH	P25701 arabidopsis
36	38	35.5	136	1 ACP3_ARATH	P25702 arabidopsis
37	38	35.5	178	1 GRPE_RICCN	Q92925 rickettsia
38	38	35.5	229	1 VATE_SPIOL	Q41396 spinacia ol
39	38	35.5	252	1 TRT3_COTJA	P06398 coturnix co
40	38	35.5	260	1 DH10_ARATH	P42759 arabidopsis
41	38	35.5	283	1 YF46_ARCFU	Q28726 archaeoglob
42	38	35.5	364	1 HEM2_STRPN	Q97r30 streptococc
43	38	35.5	411	1 MP62_LYTPI	P91753 lytechinus
44	38	35.5	498	1 IRF5_HUMAN	Q13568 homo sapien
45	38	35.5	541	1 UL21_VZVD	P09289 varicella-z

ALIGNMENTS

RESULT 1	CAS2_BOVIN	STANDARD:	PRT:	222 AA.
ID	P02663: Q9TR51:			
AC	21-JUL-1986 (Rel. 01, Created)			
DT	01-JUL-1989 (Rel. 11, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Alpha-S2 casein precursor [Contains: Casocidin-I].			
GN	CSNIS2.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID:9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=8818899; PubMed=2833669;			
RA	Stewart A.F., Bonsing J., Beattie C.W., Shah F., Willis I.M.,			
RA	Mackinlay A.G.			
RT	"Complete nucleotide sequences of bovine alpha S2- and beta-casein			
RT	cDNAs: comparisons with related sequences in other species.";			
RL	Mol. Biol. Evol. 4:231-241(1987).			
RN	[2]			
RP	SEQUENCE OF 16-222 (A ALLELE).			
RC	TISSUE=Milk;			
RX	MEDLINE=77185633; PubMed=862906;			
RA	Brignon G., Ribadeau-Dumas B., Mercier J.-C., Pelissier J.-P.,			
RA	Das B.C.;			
RT	"Complete amino acid sequence of bovine alphas2-casein.";			
RL	FEBS Lett. 76:274-279(1977).			
RN	[3]			
RP	PARTIAL SEQUENCE (D ALLELE).			
RC	TISSUE=Milk;			
RX	MEDLINE=79239837; PubMed=469044;			
RA	Grosclaude F., Joudrier P., Mahe M.-F.;			
RT	"A genetic and biochemical analysis of a polymorphism of bovine alpha			
RT	S2-casein.";			
RL	J. Dairy Res. 46:211-213(1979).			
RN	[4]			
RP	SEQUENCE OF 165-203, AND CHARACTERIZATION OF CASOCIDIN.			
RC	TISSUE=Milk;			
RX	MEDLINE=96000204; PubMed=755666;			
RA	Zucht H.-D., Raída M., Adermann K., Meagert H.-J., Forssmann W.-G.;			
RT	"Casocidin-I: a casein-alpha s2 derived peptide exhibits antibacterial			
RT	activity.";			
RL	FEBS Lett. 372:185-188(1995).			
CC	-!- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT			
CC	-!- CALCIUM PHOSPHATE.			
CC	-!- FUNCTION: CASOCIDIN-I INHIBITS THE GROWTH OF ESCHERICHIA COLI AND			
CC	STAPHYLOCOCCUS CARNOSUS.			
CC	-!- SUBCELLULAR LOCATION: Extracellular.			
CC	-!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.			
CC	-!- MASS SPECTROMETRY: MW=4870; METHOD-Electrospray; RANGE=165-203.			
CC	-!- POLYMORPHISM: AT LEAST TWO ALLELES EXIST. THE SEQUENCE OF THE A			
CC	ALLELE IS SHOWN HERE. THE D ALLELE SEQUENCE DIFFERS FROM THAT			
CC	SHOWN IN HAVING A DELETION OF NINE RESIDUES, WHICH MAY BE 49-58,			

us-09-380-738a-3.rsp

wed Feb 12 11:35:23 2003

```

OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86104467; PubMed=3002499;
RA Bolnsard M., Petrisant G.;
RT "Complete sequence of ovine alpha s2-casein messenger RNA.";
RL Biochimie 67:1043-1051(1985).
CC -!- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT
CC -!- CALCIUM PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: BELONGS TO THE ALPHA-CASEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X03238; CAA26983.1; -
CC PIR: A25070; KASHS2.
CC InterPro: IPR001588; Casein.
CC Pfam: PF00363; caseins; 2.
CC PROSITE: PS00306; CASEIN.ALPHA.BETA; 1.
CC Milk; Phosphorylation; Repeat; Signal.
KW SIGNAL 1 15
FT CHAIN 16 223 ALPHA-S2 CASEIN.
FT REPEAT 77 141
FT REPEAT 159 223
FT MOD_RES 23 23 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 24 24 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 25 25 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 72 72 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 73 73 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 74 74 PHOSPHORYLATION (POTENTIAL).
FT VARIANT 64 64 D -> N.
FT SEQUENCE 223 AA; 26332 MW; 67212935E27426D7 CRC64;
DR
DR Query Match 78.5%; Score 84; DB 1; Length 223;
DR Best Local Similarity 72.0%; Pred. No. 1.7e-06;
DR Matches 18; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 1 NANEEYSIGXXXXXAEVATEVK 25
DB 62 NADEEYSIRSSSESAEVAPEVK 86
RESULT 4
ID KDGE.DROME STANDARD; PRT; 1454 AA.
AC 009103.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Eye-specific diacylglycerol kinase (EC 2.7.1.107) (Retinal
DE degeneration A protein) (Diglyceride kinase 2) (DGK 2) (DAG kinase 2).
GN RDGA OR DGK2.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Canton-S; TISSUE=Head;
RC MEDLINE=94068563; PubMed=8248222;
RA Masai I., Okazaki A., Hosoya T., Hotta Y.;
RT "Drosophila retinal degeneration A gene encodes an eye-specific
RT diacylglycerol kinase with cysteine-rich zinc-finger motifs and
RT ankyrin repeats."

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 90:11157-11161(1993).
CC -!- FUNCTION: REQUIRED FOR THE MAINTENANCE OF THE PHOTORECEPTOR. ITS
CC ABSENCE LEADS TO RHABDOMERE DEGENERATION DUE TO DEFECTIVE
CC PHOSPHOLIPID TURNOVER.
CC -!- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-
CC diacylglycerol 3-phosphate.
CC -!- SUBCELLULAR LOCATION: Membrane-associated.
CC -!- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN ADULT EYE.
CC -!- DISEASE: MUTANTS OF THIS GENE ARE CHARACTERIZED BY HAVING
CC PHOTORECEPTOR CELLS THAT DEVELOP NORMALLY BUT DEGENERATE RAPIDLY
CC AFTER ECLOSION.
CC -!- SIMILARITY: BELONGS TO THE EUKARYOTIC DIACYLGLYCEROL KINASE
CC FAMILY.
CC -!- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC BINDING DOMAINS.
CC -!- SIMILARITY: CONTAINS 4 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: D17315; BAA04135.1; -
CC FLYBase: FBgn0003217; rdga.
CC InterPro: IPR002110; ANK.
CC Pfam: PF000756; DAGKA.
CC InterPro: IPR001206; DAGKC.
CC InterPro: IPR002219; DAG_PE-bind.
CC InterPro: IPR003622; DAG_kin_cat.
CC Pfam: PF00023; ank; 4.
CC Pfam: PF00130; DAG_PE-bind; 1.
CC Pfam: PF00609; DAGKA; 1.
CC Pfam: PF00781; DAGKC; 1.
CC ProDom: PD002939; DAGKA; 1.
CC ProDom: PD005043; DAG_kin_cat; 1.
CC SMART: SM00109; C1; 2.
CC SMART: SM00248; ANK; 2.
CC SMART: SM00045; DAGKA; 1.
CC SMART: SM00046; DAGKC; 1.
CC PROSITE: PS50088; ANK_REPEAT; 2.
CC PROSITE: PS50297; ANK_REPEAT_REGION; 1.
CC PROSITE: PS50479; DAG_PE_BIND_DOM_1; FALSE NEG.
CC PROSITE: PS50081; DAG_PE_BIND_DOM_2; FALSE NEG.
CC Transferrase; Kinase; ANK repeat; Repeat; Vision;
KW Phorbol-ester binding.
FT DOMAIN 592 642 PHORBOL-ESTER AND DAG BINDING 1.
FT DOMAIN 662 719 PHORBOL-ESTER AND DAG BINDING 2.
FT DOMAIN 807 935 CATALYTIC-A (POTENTIAL).
FT DOMAIN 961 1115 CATALYTIC-B (POTENTIAL).
FT REPEAT 1317 1346 ANK 1.
FT REPEAT 1350 1379 ANK 2.
FT REPEAT 1386 1415 ANK 3.
FT REPEAT 1419 1448 ANK 4.
FT DOMAIN 2 6 POLY-GLN.
FT DOMAIN 24 39 THR-RICH.
FT DOMAIN 110 115 POLY-SER.
FT DOMAIN 227 231 POLY-GLU.
FT DOMAIN 758 775 GLY-RICH.
FT MUTAGEN 869 869 G->D: IN MUTANT RDGA2.
FT SEQUENCE 1454 AA; 159675 MW; CE4C81099FEA16AA CRC64;
QY 1 NANEE-----YSIGXXXXXAEVATEE 23
DB 225 NADEEESGAALAEAEETTEATEE 250
Query Match 43.5%; Score 46.5; DB 1; Length 1454;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 13; Conservative 1; Mismatches 9; Indels 3; Gaps 1;
QY 1 NANEE-----YSIGXXXXXAEVATEE 23
DB 225 NADEEESGAALAEAEETTEATEE 250

```

RESULT 5

CHS5_YEAST STANDARD; PRT; 671 AA.

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Chitin biosynthesis protein CHS5 (CAL3 protein).

GN CHS5 OR CAL3 OR YLR330W OR L8543.18.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=GRF588;

RX MEDLINE=97265380; PubMed=9111317;

RA Santos B., Duran A., Valdivieso M.H.;

RT "CHS5, a gene involved in chitin synthesis and mating in

RL Saccharomyces cerevisiae";

RL Mol. Cell. Biol. 17:2485-2496(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;

RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,

RA Favell A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,

RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,

RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,

RA Miller N., Nhan M., Pauley A., Peluso D., Rifkin L., Riles L.,

RA Taich A., Trevasakis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,

RA Wilson R., Waterston R.;

RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.

RN [3]

RP FIBRONECTIN TYPE III DOMAIN.

RX MEDLINE=97148176; PubMed=8994808;

RA Bateman A., Chothia C.;

RT "Fibronectin type III domains in yeast detected by a hidden Markov

RL model";

RL Curr. Biol. 6:1544-1546(1996).

CC -!- FUNCTION: INVOLVED IN CHITIN SYNTHESIS AND ALSO REQUIRED FOR

CC MATING.

CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

EMBL; 249198; CAA89059.1; -

DR EMBL; U20618; AAB64526.1; -

DR SGD; S0004322; CHS5.

DR InterPro; IPR001357; BRCT.

DR InterPro; IPR003961; FN_III.

DR Pfam; PF00041; fn3; 1.

DR Pfam; PF00533; BRCT; 1.

DR SMART; SM00292; BRCT; 1.

DR SMART; SM00060; FN3; 1.

DR PROSITE; PS50172; BRCT; UNKNOWN_1.

DR Repeat.

FT DOMAIN 76 166 FIBRONECTIN TYPE-III.

SQ SEQUENCE 671 AA; 73638 MW; FA92741B862814C2 CRC64;

Query Match 41.1%; Score 44; DB 1; Length 671;

Best Local Similarity 55.0%; Pred. No. 14;

Matches 11; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 3 NEEEYSGXXXEEAEVATE 22

IIII I I IIII

Db 278 NEEELSYSENPVAEVADE 297

RESULT 6

YIA3_RHISP STANDARD; PRT; 88 AA.

AC P17985;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Insertion element ISR1 hypothetical 10 kba protein A3.

OS Rhizobium sp. Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Rhizobium.

OX NCBI_TaxID=391;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Class IV strains;

RX MEDLINE=89297222; PubMed=2544911;

RA Priefer U.B., Kalinowski J., Rueger B., Heumann W., Puehler A.;

RT "ISR1, a transposable DNA sequence resident in Rhizobium class IV

RL strains, shows structural characteristics of classical insertion

RL elements";

RL Plasmid 21:120-128(1989).

CC -!- SIMILARITY: BELONGS TO THE TRANSPOSASE FAMILY 8.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

EMBL; X06616; CAA29830.1; -

DR PIR; S09659; S09659.

DR InterPro; IPR002514; Transposase_8.

DR Pfam; PF01527; Transposase_8; 1.

DR KW Hypothetical protein; Transposable element.

SQ SEQUENCE 88 AA; 10024 MW; 64B7027A9FB7DB10 CRC64;

Query Match 40.2%; Score 43; DB 1; Length 88;

Best Local Similarity 55.0%; Pred. No. 2.5;

Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 5 EEYSIGXXXEEAEVATEV 24

II II I I I I I I

Db 8 EEQIGILREQEAGVATAEV 27

RESULT 7

SKI2_YEAST STANDARD; PRT; 1287 AA.

ID SKI2_YEAST

AC P35207; Q06047;

DT 01-FEB-1994 (Rel. 28, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Antiviral protein SKI2.

GN SKI2 OR YLR398C OR L8084.17.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93309467; PubMed=8321235;

RA Widner W.R., Wickner R.B.;

RT "Evidence that the SKI antiviral system of Saccharomyces cerevisiae

RL acts by blocking expression of viral mRNA";

RL Mol. Cell. Biol. 13:4331-4341(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;

RX MEDLINE=97313267; PubMed=9169871;

RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,

```

RESULT 8
NFL_BOVIN          STANDARD;          PRT;          554 AA.
ID    NFL_BOVIN
AC    P02548; P79127;
DT    21-JUL-1986 (Rel. 01, Created)
DT    01-NOV-1997 (Rel. 35, Last sequence update)
DT    16-OCT-2001 (Rel. 40, Last annotation update)
Query Match          39.3%;          Score 42;          DB 1;          Length 554;
Best Local Similarity 45.8%;          Pred. No. 24;
Matches 11;          Conservative 0;          Mismatches 13;          Indels 0;          Gaps 0;
QY    2 ANEEYSIGXXXEXAEVATEEVK 25
      | | | | | | | | | |
DB    517 AKEESGGESEAEETKAEEEKK 540

```

```

RESULT 9
ID NPRM_BACME STANDARD; PRT; 562 AA.
AC Q00891;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Bacilloisyn precursor (EC 3.4.24.28) (Neutral protease).
GN NPRM.
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14581;
RX MEDLINE=93195511; PubMed=8450307;
RA Kuhn S., Fortnagel P.;
RT "Molecular cloning and nucleotide sequence of the gene encoding a
calcium-dependent exoprotease from Bacillus megaterium ATCC
14581."
RL J. Gen. Microbiol. 139:39-47(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 319;
RX MEDLINE=94288995; PubMed=7764969;
RA Meinhardt F., Busskamp M., Wittchen K.D.;
RT "Cloning and sequencing of the leu C and npr M genes and a putative
spo IV gene from Bacillus megaterium DSM319."
RL Appl. Microbiol. Biotechnol. 41:344-351(1994).
CC -!- FUNCTION: THERMOSTABLE EXTRACELLULAR ZINC METALLOPROTEASE.
CC -!- CATALYTIC ACTIVITY: Similar, but not identical, to that of
thermolysin.
CC -!- COFACTOR: BINDS 1 ZINC ION AND FOUR CALCIUM IONS (BY
SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X61380; CAA43654.1; -
CC EMBL; X75070; CAA52964.1; -
CC HSSP; P05806; INPC.
CC MEROPS; M04.015; -
CC InterPro; IPR005075; Pep_M4_propep.
CC InterPro; IPR001570; Peptidase_M4.
CC InterPro; IPR000130; Zn_Mtpeptidse.
CC Pfam; PF01447; Peptidase_M4; 1.
CC Pfam; PF02868; Peptidase_M4_C; 1.
CC Pfam; PF03413; Pep_M4_propep; 1.
CC PRINTS; PR00730; THERMOLYSIN.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Calcium; Zymogen; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 245 ACTIVATION PEPTIDE (POTENTIAL).
FT CHAIN 246 562 BACILLOLYSIN.
FT METAL 388 388 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 389 389 BY SIMILARITY.
FT METAL 392 392 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 412 412 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 477 477 PROTON DONOR (BY SIMILARITY).
FT CONFLICT 83 83 S -> N (IN REF. 2).
FT CONFLICT 186 186 A -> G (IN REF. 2).
FT CONFLICT 251 251 T -> A (IN REF. 2).
FT CONFLICT 302 302 A -> T (IN REF. 2).
FT CONFLICT 344 344 R -> A (IN REF. 2).

```

```

FT CONFLICT 394 394 L -> V (IN REF. 2).
SQ SEQUENCE 562 AA; 60948 MW; 17203441C7F6AAB7 CRC64;

Query Match 39.3%; Score 42; DB 1; Length 562;
Best Local Similarity 32.0%; Pred. No. 25;
Matches 8; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 NANEEEYSIGXXXEEAEVATEEVK 25
DB 71 NTNKDKYKLGNSAQNSFKVTEVVK 95
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |

RESULT 10
ARP8_YEAST STANDARD; PRT; 881 AA.
AC Q12386;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Actin-like protein ARP8.
GN ARP8 OR YOR141C OR YOR3348C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97344368; PubMed=9200815;
RA Voss H., Benes V., Andrade M.A., Valencia A., Rechmann S., Teodoru C.,
RA Schwager C., Paces V., Sander C., Ansorge W.;
RT "DNA sequencing and analysis of 130 kb from yeast chromosome XV."
RL Yeast 13:655-672(1997).
RN [2]
RP GENE NAME.
RX MEDLINE=97435478; PubMed=9290209;
RA Poch O., Winsor B.;
RT "Who's who among the Saccharomyces cerevisiae actin-related proteins?
A classification and nomenclature proposal for a large family."
RL Yeast 13:1053-1058(1997).
CC -!- SIMILARITY: BELONGS TO THE ACTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X94335; CAA64058.1; -
CC EMBL; Z75049; CAA99341.1; -
CC SGD; S0005667; ARP8.
CC InterPro; IPR004000; Actin_like.
CC Pfam; PF00022; actin; 2.
CC SMART; SM00268; ACTIN; 1.
KW Structural protein; Cytoskeleton.
FT DOMAIN 22 27 POLY-ASP.
SQ SEQUENCE 881 AA; 100208 MW; 8174851B6B077A19 CRC64;

Query Match 39.3%; Score 42; DB 1; Length 881;
Best Local Similarity 36.0%; Pred. No. 39;
Matches 9; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 1 NANEEEYSIGXXXEEAEVATEEVK 25
DB 108 NANENENELGSSRDKRAPPVQTSK 132
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |

RESULT 11
YB4F_SCHPO STANDARD; PRT; 516 AA.
ID YB4F_SCHPO
AC O14360;
DT 15-JUL-1998 (Rel. 36, Created)

```


DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C30D10.15 in chromosome II.
 GN SPBC30D10.15.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman C.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Ocell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.W.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga K.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shipakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RA "The genome sequence of Schizosaccharomycetes pombe."
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: TO YEAST YNL124W.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z97992; CAB10810.1; -.
 DR Hypothetical protein.
 KW SEQUENCE 516 AA; 57639 MW; 22597CE8E5AF9B5F CRC64;
 SQ
 Query Match 38.8%; Score 41.5; DB 1; Length 516;
 Best Local Similarity 36.0%; Pred. No. 27;
 Matches 9; Conservative 8; Mismatches 3; Indels 5; Gaps 1;
 QY 1 NANEESYSGXXEAEVATEEVK 25
 | :|:|:| :|:|:|:|
 DB 359 NPEQEFSS-----DDEAEVAAKOLK 378
 RESULT 12
 ID NFL_PIG STANDARD; PRT; 548 AA.
 AC P02547;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 DE Neurofilament light polypeptide (NF-L).
 GN NEFL.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85154583; PubMed=3920075;
 RA Geisler N., Plessmann U., Weber K.;
 RT "The complete amino acid sequence of the major mammalian
 RT neurofilament protein (NF-L).";
 RL FEBS Lett. 182:475-478(1985).
 RN [2]
 RP SEQUENCE OF 1-82 AND 278-548.
 RA Geisler N., Kaufmann E., Fischer S., Plessmann U., Weber K.;
 RT "Neurofilament architecture combines structural principles of
 RT intermediate filaments with carboxy-terminal extensions increasing
 RT in size between triplet proteins.";
 RL EMBO J. 2:1295-1302(1983).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 DR PIR; A02963; OPFGL.
 DR InterPro; IPR001664; IF.
 DR Pfam; PF00038; filament; 1.
 DR PROSITE; PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neurone; Glycoprotein.
 FT DOMAIN 1 91
 FT DOMAIN 92 395
 FT DOMAIN 396 548
 FT DOMAIN 549 123
 FT DOMAIN 124 136
 FT DOMAIN 137 232
 FT DOMAIN 233 251
 FT DOMAIN 252 270
 FT DOMAIN 271 279
 FT DOMAIN 280 395
 FT DOMAIN 396 442
 FT DOMAIN 443 548
 FT CARBOHYD 20 26
 FT CARBOHYD 26 20
 FT SITE 380 390
 FT UNSURE 322 322
 FT SEQUENCE 548 AA; 61940 MW; 83044813637AC739 CRC64;
 SQ
 Query Match 38.3%; Score 41; DB 1; Length 548;
 Best Local Similarity 45.8%; Pred. No. 35;
 Matches 11; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 2 ANEEYSIGXXEAEVATEEVK 25
 | :|:|:| :|:|:|:|
 DB 511 AKEEGEGEGQGEETKEAEVEEK 534
 RESULT 13
 ID Y012_CAEEL STANDARD; PRT; 696 AA.
 AC P34668;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative ATP-dependent RNA helicase ZK686.2 in chromosome III.
 DE ZK686.2.
 GN Caenorhabditis elegans.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditoidea;

```

RA Kappeler S., Farah Z., Puhan Z.;
RT "Sequence analysis of Camelus dromedarius milk caseins.";
RL J. Dairy Res. 65:209-222(1998).
CC -I- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT
CC CALCIUM PHOSPHATE.
CC -I- SUBCELLULAR LOCATION: Extracellular.
CC -I- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -I- SIMILARITY: BELONGS TO THE ALPHA-CASEIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ012629; CAAL0078.1; -.
DR InterPro; IPR001588; Casein.
DR Pfam; PF00363; Caseins; 2.
DR PROSITE; PS00306; CASEIN_ALPHA_BETA; FALSE_NEG.
KW Milk; Phosphorylation; Signal.
FT SIGNAL 1 15
FT CHAIN 16 193
FT -----
SQ SEQUENCE 193 AA; 22964 MW; 2843256F8FD2ED13 CRC64;

Query Match 37.9%; Score 40.5; DB 1; Length 193;
Best Local Similarity 52.4%; Pred. No. 14;
Matches 11; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 3 NEEYSIGXXKXEEAEVATEE 23
Db 118 NTEQLSI---SEESTEVPTPE 135

RESULT 15
ACPI_ARATH
ID ACPI_ARATH STANDARD; PRT; 137 AA.
AC P11829;
DC 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Acyl carrier protein 1, chloroplast precursor (ACP).
GN ACP-1 OR AT3G05020 OR T9J14.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
OC [1]
RP SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RP MEDLINE=89160348; PubMed=2922299;
RX Post-Bettenmiller M.A.; Hlousek-Radojcic A.; Ohlroage J.B.;
RT "DNA sequence of a genomic clone encoding an Arabidopsis acyl carrier
RT protein (ACP).";
RL Nucleic Acids Res. 17:1777-1777(1989).
RL [2]
RP REVISIONS.
RP Post-Bettenmiller M.A.;
RL Submitted (FEB-1989) to the EMBL/GenBank/DBJ databases.
RL [3]
RP SEQUENCE FROM N.A.
RP STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M.; Lemcke K.; Rieger M.; Ansoerge W.; Unseld M.,
RA Farlmann B.; Valté G.; Bloecker H.; Perez-Alonso M.; Obermaier B.,
RA Delseny M.; Botry M.; Grivell L.A.; Mache R.; Puigdomenech P.,
RA De Simone V.; Choise N.; Artiguenave F.; Robert C.; Brottier P.,
RA Wincker P.; Cattolico L.; Weissenbach J.; Saurin W.; Quétier F.,
RA Schwaebler M.; Mueller-Auer S.; Gabel C.; Fuchs M.; Benes V.,
RA Wurmbach E.; Drzonek H.; Erfle H.; Jordan N.; Bangert S.,
RA Wiedelmann R.; Kranz H.; Voss H.; Holland R.; Brandt P.; Nyakatura G.,

```

Wed Feb 12 11:35:23 2003

RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,
RA Reichelt J., Scharfe M., Schoen O., Barques M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujil C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pal G., Millischer J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nieman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Idesawa K., Kawashima K., Kishida Y.,
RA Sasamoto S., Kimura T., Matsumoto M., Matsuno A., Muraki A.,
RA Kiyokawa C., Kohara M., Nakazaki N., Shimpou S., Takeuchi C., Wada T.,
RA Nakayama S., Yamada M., Yasuda M., Tabata S.,
RA Watanabe A.,
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RL thaliana.";
RL Nature 408:820-822(2000).
RN [4]
RN CHARACTERIZATION.
RP Hlousek-Radojcic A., Post-Beittemiller D., Ohlrogge J.B.;
RA "Expression of constitutive and tissue-specific acyl carrier protein
RT isoforms in Arabidopsis.";
RT Plant Physiol. 98:206-214(1992).
RL -!- FUNCTION: ACYL CARRIER PROTEIN IS A KEY COMPONENT IN DE NOVO FATTY
CC ACID BIOSYNTHESIS. IT IS A SMALL ACIDIC PROTEIN WITH A 4'-PHOSPHO-
CC PANTHETINE PROSTHETIC GROUP, ATTACHED THROUGH A SERINE, TO WHICH
CC GROWING FATTY ACID CHAINS ARE COVALENTLY BOUND.
CC -!- SIMILARITY: CONTAINS 1 ACYL CARRIER DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X13708; CAA31991.1; -;
CC EMBL; AC009465; AAG51406.1; -;
CC PIR; S03267; S03267.
CC HSP; P02901; IACP.
CC InterPro: IPR003231; Acyl_carrier.
CC InterPro: IPR003880; Pantne_attach.
CC Pfam; PF00550; pp-binding; 1
CC ProDom; PD000887; Acyl_carrier; 1.
CC TIGRFAMs; TIGR00517; acyl_carrier; 1.
CC PROSITE; PS00012; PHOSPHOPANTHETINE; 1.
CC PROSITE; PS00075; ACP_DOMAIN; 1.
CC Fatty acid biosynthesis; phosphopantetheine; Chloroplast;
CC Transit peptide; Multigene family.
CC TRANSIT 1 54 CHLOROPLAST.
CC CHAIN 55 137 ACYL CARRIER PROTEIN 1.
CC BINDING 93 93 PHOSPHOPANTHETINE (BY SIMILARITY).
CC SEQUENCE 137 AA; 15055 MW; 6AA3431A78640C6D CRC64;
CC
CC Query Match 37.4%; Score 40; DB 1; Length 137;
CC Best Local Similarity 40.0%; Pred. No. 12;
CC Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 4 EEEYIGXXXXXAEVATEE 23
Db 104 EEEFNQMAEERAKIATVE 123

Search completed: February 11, 2003, 18:17:46
Job time : 12.0825 secs


```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on:      February 11, 2003, 18:13:41 ; Search time 28.0928 Seconds
              (without alignments)
              183.363 Million cell updates/sec

Title:       US-09-380-738A-3
Perfect score: 107
Sequence:    1 NANEEYSIGXXXEXAEVATEEVK 25

Scoring table:  BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched:      671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

0
5
5
5
2
2
0

```

4:  sp_human.*
5:  sp_invertebrate.*
6:  sp_mammal.*
7:  sp_mmc.*
8:  sp_organelle.*
9:  sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

```

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	103	96.3	210	6	O62825		O62825 bubalus bub
2	88	82.2	124	6	O9MYU7		O9MYU7 capra hircu
3	88	82.2	223	6	O9TTQ7		O9TTQ7 capra hircu
4	88	82.2	223	6	O9MYU6		O9MYU6 capra hircu
5	88	82.2	223	6	O9GK07		O9GK07 capra hircu
6	88	82.2	914	5	O8WR09		O8WR09 drosophila
7	50	46.7	914	5	O8WR09		O8WR09 drosophila
8	50	46.7	997	5	O8STB9		O8STB9 drosophila
9	50	46.7	1179	5	O8WR08		O8WR08 drosophila
10	48	44.9	1277	5	O9V717		O9V717 drosophila
11	47	43.9	414	5	O8STL0		O8STL0 encephalito
12	46.5	43.5	1280	10	O9SFD2		O9SFD2 arabidopsis
13	46.5	43.5	1417	5	O9W3A4		O9W3A4 drosophila
14	46	43.0	526	4	O8SY47		O8SY47 drosophila
15	46	43.0	528	4	O9UP59		O9UP59 homo sapien
16	46	43.0	562	4	O95228		O95228 homo sapien
17	46	43.0	562	4	O12820		O12820 homo sapien

```

DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Alpha s2-casein.
GN CSN1S2.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RA Lagonigro R., Pilla F., Matassino D., Zullo A.;
RT "A new allele of goat alpha s2-casein gene."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ289715; CAB94235.1;
DR InterPro; IPR001588; Casein.
DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
SQ SEQUENCE 124 AA; 14533 MW; C363E536CC17B5F2 CRC64;

Query Match      82.2%; Score 88; DB 6; Length 124;
Best Local Similarity 72.0%; Pred. No. 1.5e-07;
Matches 18; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 NANEESYSGXXXEEAEVATEEVK 25
Db 62 NANEESYSSRSSEAEVATEEVK 86

RESULT 3
Q9TTQ7
ID Q9TTQ7 PRELIMINARY; PRT; 223 AA.
AC Q9TTQ7
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Alpha s2-casein.
GN CSN1S2.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RA Veltri C.C., Pilla F., Lagonigro R.R.;
RT "A new allele of goat alpha s2-casein."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ249995; CAB59920.1;
DR InterPro; IPR001588; Casein.
DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
SQ SEQUENCE 223 AA; 26433 MW; CE9F4DC8D7688293 CRC64;

Query Match      82.2%; Score 88; DB 6; Length 223;
Best Local Similarity 72.0%; Pred. No. 3e-07;
Matches 18; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 NANEESYSGXXXEEAEVATEEVK 25
Db 62 NANEESYSSRSSEAEVATEEVK 86

RESULT 4
Q9MYU6
ID Q9MYU6 PRELIMINARY; PRT; 223 AA.
AC Q9MYU6
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Alpha s2-casein.
GN CSN1S2.
OS Capra hircus (Goat).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RA Lagonigro R., Pilla F., Matassino D., Zullo A.;
RT "A new allele of goat alpha s2-casein gene."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ289716; CAB94236.1;
DR InterPro; IPR001588; Casein.
DR Pfam; PF00363; caseins; 2.
DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
SQ SEQUENCE 223 AA; 26403 MW; 0E1FE83F24DA85E2 CRC64;

Query Match      82.2%; Score 88; DB 6; Length 223;
Best Local Similarity 72.0%; Pred. No. 3e-07;
Matches 18; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 NANEESYSGXXXEEAEVATEEVK 25
Db 62 NANEESYSSRSSEAEVATEEVK 86

RESULT 5
Q9GK07
ID Q9GK07 PRELIMINARY; PRT; 223 AA.
AC Q9GK07
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Alpha s2-casein.
GN CSN1S2.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RA Veltri C., Pilla F., Lagonigro R.;
RT "A new allele of alpha s2-casein."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ297311; CAC21704.2;
DR EMBL; AJ242728; CAC21704.2; JOINED.
DR EMBL; AJ297312; CAC21704.2; JOINED.
DR EMBL; AJ297313; CAC21704.2; JOINED.
DR EMBL; AJ297314; CAC21704.2; JOINED.
DR EMBL; AJ242527; CAC21704.2; JOINED.
DR EMBL; AJ297315; CAC21704.2; JOINED.
DR EMBL; AJ297316; CAC21704.2; JOINED.
DR EMBL; AJ242526; CAC21704.2; JOINED.
DR EMBL; AJ242528; CAC21704.2; JOINED.
DR InterPro; IPR001588; Casein.
DR Pfam; PF00363; caseins; 2.
DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 2.
SQ SEQUENCE 223 AA; 26432 MW; CE9765E8D7688C9D CRC64;

Query Match      82.2%; Score 88; DB 6; Length 223;
Best Local Similarity 72.0%; Pred. No. 3e-07;
Matches 18; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 NANEESYSGXXXEEAEVATEEVK 25
Db 62 NANEESYSSRSSEAEVATEEVK 86

RESULT 6
Q8WR09
ID Q8WR09 PRELIMINARY; PRT; 914 AA.
AC Q8WR09;

```

DE Dystroglycan type I.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeuchi K.-I., Yamaguchi A., Kaneda M., Aizu M., Umeda M.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF461100; AAL66367.1; -
SQ SEQUENCE 914 AA; 102660 MW; E255B90EA9DDA76C CRC64;

Query Match 46.7%; Score 50; DB 5; Length 914;
Best Local Similarity 47.6%; Pred. No. 6;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 ANEEYSIGXXXEAEVATE 22
|:::| | |||||
Db 207 ADEDDYDGDDEVAEPSTE 227

RESULT 7
Q8STB9 PRELIMINARY; PRT; 997 AA.

ID Q8STB9
AC Q8STB9; (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Dystroglycan type II (GH09323p).
GN DG.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeuchi K.-I., Yamaguchi A., Kaneda M., Aizu M., Umeda M.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY069076; AAL39221.1; -
SQ SEQUENCE 997 AA; 110752 MW; 7F3B5E7947CD9A24 CRC64;

Query Match 46.7%; Score 50; DB 5; Length 997;
Best Local Similarity 47.6%; Pred. No. 6.6;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 ANEEYSIGXXXEAEVATE 22
|:::| | |||||
Db 207 ADEDDYDGDDEVAEPSTE 227

RESULT 8
Q8WR08 PRELIMINARY; PRT; 1179 AA.

ID Q8WR08
AC Q8WR08; (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

1

Adams M.D., Celniker S.E., Holl R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Andrews-Pfannkoch C., Baldwin D.,
Abril J.F., Agbayani A., An H.-J., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Ballew R.M., Basu A., Bonos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Beeson K.Y., Benos P.V., Berman B.P., Borkov D., Botchan M.R., Bouck J.,
Burtis K.C., Busam D.A., Butler H., Brokstein P., Brotter P.,
Cherry J.M., Cawley S., Dahlke C., Cadieu E., Center A., Chandra I.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
A Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
A Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
A Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
A Kimmel B.E., Kodira C.D., Kraft C., Li J., Li Z., Liang Y., Lin X.,
A Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
A Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
A Mervulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
A Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
A Nelson D.R., Nelson K.A., Nixon K., Nusskern D.J., Pacle J.M.,
A Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri Y., Resse M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL: AE003445; AAF46430.1; -.
DR FlyBase: FBgn0003217; rdga.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000756; DAGKa.
DR InterPro: IPR001206; DAGKc.
DR InterPro: IPR003622; DAG_kin_cat.
DR InterPro: IPR002219; DAG_PE-bind.
DR Pfam: PF00023; ank; 4.
DR Pfam: PF00609; DAGKa; 1.
DR Pfam: PF00781; DAGKc; 1.
DR ProDom: PD002939; DAGKa; 1.
DR ProDom: PD005043; DAG_kin_cat; 1.
DR SMART: SM00248; ANK; 2.
DR SMART: SM00109; Cl; 2.
DR SMART: SM00045; DAGKa; 1.
DR SMART: SM00046; DAGKc; 1.
DR PROSITE: PS50088; ANK_REPEAT; 2.
DR PROSITE: PS50297; ANK_REPEAT; 1.
DR ANK repeat; Repeat.
KW ANK repeat; Repeat.
SQ SEQUENCE 1417 AA; 155576 MW; DBEC5CA6517711AB CRC64;

Query Match 43.5%; Score 46.5; DB 5; Length 1417;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 13; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 1 NANEEE---YSIGXXXEXAEVATEE 23
DB 225 NAEEEEGSAIADEAEETTEATEE 250

RESULT 13
Q8SY47 ID Q8SY47 PRELIMINARY; PRT; 1457 AA.
AC Q8SY47
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE GH23785p.
GN RDGA.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guatini H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL: AY075349; AAL68208.1; -.
DR EMBL: 1457 AA; 160141 MW; 8F4E33E9C1B665D1 CRC64;
SQ SEQUENCE 1457 AA; 160141 MW; 8F4E33E9C1B665D1 CRC64;

Query Match 43.5%; Score 46.5; DB 5; Length 1457;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 13; Conservative 1; Mismatches 9; Indels 3; Gaps 1;

QY 1 NANEEE---YSIGXXXEXAEVATEE 23

DB 225 NAEEEEGSAIADEAEETTEATEE 250

RESULT 14
Q9UP59 ID Q9UP59 PRELIMINARY; PRT; 526 AA.
AC Q9UP59
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE PITSLRE protein kinase alpha SV4 isoform.
GN CDC2L1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98424414; PubMed=9750192;
RA Gururajan R., Lahti J.M., Grenet J., Easton J., Gruber I.,
RA Ambros P.F., Kidd V.J.;
RT "Duplication of a genomic region containing the Cdc2L1-2 and MMP21-22
RT genes on human chromosome 1p36.3 and their linkage to D122.";
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AF067513; AAC72078.1; -.
DR HSSP: P24941; 1B38.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW SEQUENCE 526 AA; 59272 MW; 3DF4854ED82B81B CRC64;
SQ SEQUENCE 526 AA; 59272 MW; 3DF4854ED82B81B CRC64;

Query Match 43.0%; Score 46; DB 4; Length 526;
Best Local Similarity 47.6%; Pred. No. 16;
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 EEEYSIGXXXEXAEVATEEV 24
DB 63 EEEETGCSNSEEAEQSAEEV 83

RESULT 15
Q95228 ID Q95228 PRELIMINARY; PRT; 528 AA.
AC Q95228
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE PITSLRE protein kinase beta SV13 isoform.
GN CDC2L2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98424414; PubMed=9750192;
RA Gururajan R., Lahti J.M., Grenet J., Easton J., Gruber I.,
RA Ambros P.F., Kidd V.J.;
RT "Duplication of a genomic region containing the Cdc2L1-2 and MMP21-22
RT genes on human chromosome 1p36.3 and their linkage to D122.";
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AF067525; AAC72090.1; -.
DR HSSP: P24941; 1B38.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
SQ SEQUENCE 528 AA; 59272 MW; 3DF4854ED82B81B CRC64;

DR Pfam: PF00069; pkinase; 1.
 DR PRODom: PD000001; Euk_pkinase; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 528 AA; 61958 MW; AE2540CF3A8E7890 CRC64;

Query Match 43.0%; Score 46; DB 4; Length 528;
 Best Local Similarity 47.6%; Pred. NO. 16;
 Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 EEYSIGXXXEEAEVATEEV 24
 ||| | ||:|: |||
 Db 256 EEEETGSGNSEAEQSAEEV 276

Search completed: February 11, 2003, 18:19:44
 Job time : 30.0928 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:06:55 ; Search time 30.5258 Seconds
(without alignments)
91.669 Million cell updates/sec

Title: US-09-380-738A-4
Perfect score: 91
Sequence: 1 KNPMHVHXXEESIIHQETVK 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
A.Geneseq_101002.*			
1:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*		
2:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*		
3:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*		
4:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*		
5:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*		
6:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*		
7:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*		
8:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*		
9:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*		
10:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*		
11:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*		
12:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*		
13:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*		
14:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*		
15:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*		
16:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*		
17:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*		
18:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*		
19:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*		
20:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*		
21:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*		
22:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*		
23:	/SID22/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	87	95.6	21	AA32934	Casein phosphopeptide
2	87	95.6	21	AA32934	Sequence of casein
3	87	95.6	21	AA32934	Bos alpha-S2-casein
4	87	95.6	24	AA32934	Anti-acid peptide.
5	87	95.6	22	AA32934	Bovine alpha-S2 ca
6	87	95.6	22	AA32934	Bovine alpha-S2 ca
7	84	92.3	21	AA32934	Anticariogenic pho
8	82	90.1	20	AA32934	Phosphopeptide 3.
9	82	90.1	20	AA32934	Phosphopeptide #3
10	82	90.1	20	AA32934	Anticariogenic pho

11	82	90.1	20	14	AA32933	Casein phosphopept
12	82	90.1	20	14	AA32933	Phosphopeptide der
13	82	90.1	20	14	AA32933	Sequence of casein
14	82	90.1	20	15	AA32933	Sodium caseinate t
15	82	90.1	20	16	AA32933	Phosphopeptide r3.
16	82	90.1	20	21	AA32933	Phosphopeptide r3.
17	59.5	65.4	23	23	AA32933	Pig alpha-S2 casei
18	54.5	59.9	22	23	AA32933	Alpha-S2 casein pr
19	53.5	58.8	22	23	AA32933	Goat alpha-S2 case
20	53.5	58.8	22	23	AA32933	Capra hircus alpha
21	43	47.3	408	22	AA32933	Novel human diagno
22	42	46.2	566	17	AA32933	B.t. neutral prote
23	39.5	43.4	223	23	AA32933	Alpha-S2 casein pr
24	39.5	43.4	295	20	AA32933	Chlamydia pneumoni
25	39	42.9	2206	21	AA32933	Plasmodium falcipa
26	38	41.8	132	22	AA32933	Human nervous syst
27	38	41.8	242	22	AA32933	Drosophila melanog
28	38	41.8	312	23	AA32933	Human serine-threo
29	38	41.8	337	23	AA32933	Human serine-threo
30	38	41.8	436	22	AA32933	Human NIM1 kinase
31	38	41.8	436	22	AA32933	Human protein kina
32	38	41.8	436	22	AA32933	Novel protein kina
33	37	40.7	92	17	AA32933	HCMV Toledo strain
34	37	40.7	121	22	AA32933	Escherichia coli p
35	37	40.7	149	21	AA32933	Arabidopsis thalia
36	37	40.7	149	21	AA32933	Arabidopsis thalia
37	37	40.7	353	22	AA32933	Human papillomavir
38	37	40.7	477	22	AA32933	Drosophila melanog
39	37	40.7	566	16	AA32933	Protease NprL. La
40	37	40.7	742	22	AA32933	Novel human diagno
41	37	40.7	875	22	AA32933	Drosophila melanog
42	37	40.7	879	22	AA32933	Drosophila melanog
43	37	40.7	938	22	AA32933	Novel human diagno
44	37	40.7	1355	22	AA32933	Novel human diagno
45	37	40.7				

ALIGNMENTS

RESULT 1
AA32934
ID AA32934 standard; Protein; 21 AA.
XX
AC AA32934;
XX
DT 02-JUL-1993 (first entry)
XX
DE Casein phosphopeptide #8.
XX
KW Casein; phosphopeptide; dental calculus; salts; alkaline metal;
KW alkaline earth metal; Zn/phosphopeptide complex; aggregate;
KW anti-carries; anti-gingivitis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 8 /label= Phosphoserine
FT Modified-site 9 /label= Phosphoserine
FT Modified-site 10 /label= Phosphoserine
FT Modified-site 16 /label= Phosphoserine
FT Modified-site 16 /label= Phosphoserine
XX WO9303707-A.
XX 04-MAR-1993.
XX 21-AUG-1992;
XX 22-AUG-1991;
XX 91US-0748344.

Wed Feb 12 11:35:23 2003

CC supplements to increase calcium bioavailability and to help prevent
 CC diseases associated with calcium deficiencies. They are particularly
 CC useful for treatment or prevention of dental caries, calcium
 CC malabsorption and bone diseases such as osteoporosis and osteomalacia.
 CC The compositions are useful in humans and in veterinary medicine in
 CC domestic animals such as cattle, sheep, horses and companion animals e.g.
 CC cats and dogs as well as zoo animals. The present sequence represents a
 CC phosphopeptide component of a specifically claimed complex.
 XX
 XX Sequence 21 AA;

Query Match 95.6%; Score 87; DB 19; Length 21;
 Best Local Similarity 81.0%; Pred. No. 3.7e-09;
 Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXEESIIIXQETK 21
 DB 1 KNTMEHVSSESIIISQETK 21

RESULT 4
 AAR42890
 ID AAR42890 standard; peptide; 24 AA.

AC AAR42890;
 DT 16-MAY-1994 (first entry)

XX Anti-acid peptide.

XX anti-acid; antiulcer; food; drink; side-effect; whey milk;
 XX whey protein condensate; WPC; skimmed milk.

XX Bos taurus.

XX Key Location/Qualifiers
 XX Modified-site 8 /note= "residue is phosphorylated"
 XX Modified-site 9 /note= "residue is phosphorylated"
 XX Modified-site 10 /note= "residue is phosphorylated"
 XX Modified-site 16 /note= "residue is phosphorylated"

XX JP05262793-A.

XX 12-OCT-1993.

XX 18-MAR-1992; 92JP-0092163.

XX 18-MAR-1992; 92JP-0092163.

XX (SNOW) SNOW BRAND MILK PROD CO LTD.

XX WPI; 1993-357195/45.

XX New aminoacid peptide(s) with no side-effects - useful as antacid
 XX drugs and antiulcer drugs and used in foods and drinks

XX Claim 1; Page 2; 7pp; Japanese.

XX The peptide can be used as an anti-acid / anti-ulcer drug. It can
 XX also be used in food or drink production and has no side effects.
 XX The peptide is prepared from milk whey, or from whey protein
 XX condensate.

XX Sequence 24 AA;

Query Match 95.6%; Score 87; DB 14; Length 24;
 Best Local Similarity 81.0%; Pred. No. 4.3e-09;
 Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXEESIIIXQETK 21
 DB 1 KNTMEHVSSESIIISQETK 21

RESULT 5
 AAW32220
 ID AAW32220 standard; protein; 222 AA.

AC AAW32220;
 DT 03-FEB-1998 (first entry)

XX Bovine alpha-S2 casein precursor.

XX Bovine alpha-S2 casein precursor; growth promoting; mitogenic assay;

XX platelet-derived growth factor; insulin-like derived growth factor;

XX Bos taurus.

XX WO9716460-A1.

XX 09-MAY-1997.

XX 31-OCT-1996; 96WO-GB02658.

XX 31-OCT-1995; 95GB-0022302.

XX (OYLI-) UNIV LIVERPOOL.

XX Liu Q, Smith JA, Wilkinson MC;

XX WPI; 1997-272048/24.

XX Manufacture of medicament or foodstuff for promoting growth - using

XX peptide(s) with a sequence identical to the C-terminal end of an

XX alpha-S2 casein precursor

XX Disclosure; Page 3; 33pp; English.

XX The present sequence represents bovine alpha-S2 casein precursor.

XX Peptides having an amino acid sequence which is substantially identical

XX to the C-terminal end of an alpha-S2 casein precursor, are used for

XX manufacture of a medicament or foodstuff for promoting growth in humans

XX or animals.

XX Sequence 222 AA;

Query Match 95.6%; Score 87; DB 18; Length 222;

Best Local Similarity 81.0%; Pred. No. 4.8e-08;

Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXEESIIIXQETK 21

DB 16 KNTMEHVSSESIIISQETK 36

RESULT 6

AAEL7468

ID AAEL7468 standard; Protein; 222 AA.

XX AAEL7468;

XX 22-APR-2002 (first entry)

XX Bovine alpha-S2 casein precursor protein.

XX Bovine; alpha-S2 casein; dermatological; antiinflammatory; fibroblast;

XX collagen; keratinocyte; skin regeneration; medicament; toothpaste; aging;

XX chewing gum; cosmetic; wrinkling; periodontal disease.

XX Bos sp.

PN WO200202133-A2.
 PD 10-JAN-2002.
 XX 13-JUN-2001; 2001WO-GB02601.
 PF 30-JUN-2000; 2000GB-0016189.
 XX (PEPS-) PEPSYN LTD.
 PA Smith JA;
 XX WPI; 2002-154690/20.
 DR Use of peptide or its derivative containing an amino acid sequence in
 PT alpha-S2 casein precursor in the manufacture of a medicament for
 PT alleviating or preventing periodontal disease and an effect of aging in
 PT skin -
 XX Claim 8; Page 6; 27pp; English.
 PS The invention relates to a composition comprising a peptide or its
 CC derivative. The peptide contains an amino acid sequence from alpha-S2
 CC casein precursor. The peptides stimulate the growth of fibroblasts, and
 CC thus the synthesis and secretion of collagen. The peptides also stimulate
 CC the growth of keratinocytes, which aid in the formation and regeneration
 CC of skin surface. The peptide is useful in the manufacture of a medicament
 CC in the form of a toothpaste or a chewing gum, for alleviating or
 CC preventing periodontal disease and a medicament in the form of a cosmetic
 CC composition for alleviating or preventing an effect of aging,
 CC particularly wrinkling of the skin. The present sequence is bovine
 CC alpha-S2 casein precursor protein.
 XX
 SQ Sequence 222 AA;
 Query Match 95.6%; Score 87; DB 23; Length 222;
 Best Local Similarity 81.0%; Pred. No. 4.8e-08;
 Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 KNTMEHVXXXESIIQETVK 21
 Db 16 KNTMEHVSSSESIISQETVK 36
 RESULT 7
 AAR28432
 ID AAR28432 standard; peptide; 21 AA.
 XX AAR28432;
 AC 19-MAR-1993 (first entry)
 DT Anticariogenic phosphopeptide.
 DE Casein; metal ion; dietetic; purification; growth medium; dietary
 KW supplement; fertiliser.
 KW Key Location/Qualifiers
 XX Modified-site 8
 FT /note= "post-translationally phosphorylated serine"
 FT Modified-site 9
 FT /note= "post-translationally phosphorylated serine"
 FT Modified-site 10
 FT /note= "post-translationally phosphorylated serine"
 FT Modified-site 16
 FT /note= "post-translationally phosphorylated serine"
 XX WO9218526-A.
 PN 29-OCT-1992.
 XX 16-APR-1992; 92WO-AU00175.
 PD
 PF
 XX

PR 19-APR-1991; 91AU-0005706.
 XX (UYME-) UNIV MELBOURNE.
 PA (VICT-) VICTORIAN DAIRY IND AUTHORITY.
 XX Reynolds EC;
 XX WPI; 1992-382039/46.
 DR Selected phospho-peptide(s) prodn. having anticariogenic
 XX activities etc. - comprises digesting soluble monovalent cation
 PT salt of casein in soln.; introducing di- or trivalent metal ion
 PT and filtering through filter having mol. wt. exclusion limit
 XX Claim 11; Page 14; 18pp; English.
 PS The peptide may be prepd. by completely digesting casein in soln.
 CC with a proteolytic enzyme, adding mineral acid to the soln. to
 CC adjust the pH to 4.7, removing any precipitate, adding calcium
 CC chloride to cause aggregation of the peptides in soln. and
 CC separating the aggregated phosphopeptides. This method allows prodn.
 CC of the phosphopeptide by industrial methods. The phosphopeptide
 CC has anticariogenic activity and may be used as a dietetic. The
 CC peptide may be used in a microbiological growth medium, as a dietary
 CC supplement or as a fertiliser.
 CC See also AAR28425-33.
 XX
 SQ Sequence 21 AA;
 Query Match 92.3%; Score 84; DB 13; Length 21;
 Best Local Similarity 76.2%; Pred. No. 1.3e-08;
 Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 KNTMEHVXXXESIIQETVK 21
 Db 1 KNTMEHVSSSESIISQETVK 21
 RESULT 8
 AAR71322
 ID AAR71322 standard; peptide; 20 AA.
 XX AAR71322;
 AC 07-MAY-1991 (first entry)
 DT Phosphopeptide 3.
 DE Caries; gingivitis; periodontal disease; osteoporosis; osteomalacia.
 KW Synthetic.
 XX Key Location/Qualifiers
 XX Modified-site 7 /label= phosphoserine
 FT Modified-site 8 /label= phosphoserine
 FT Modified-site 9 /label= phosphoserine
 FT Modified-site 15 /label= phosphoserine
 FT Modified-site 15 /label= phosphoserine
 XX WO8707616-A.
 PN 17-DEC-1987.
 PD 12-JUN-1987; 87WO-AU00172.
 XX 12-JUN-1986; 86AU-0006385.
 PR (VICT-) VICTORIA DAIRY INDUSTRY AUTHORITY.
 PA (UYME-) UNIVERSITY OF MELBOURNE.
 XX (REYN/) EC REYNOLDS.
 PA

XX Reynolds EC;
XX WPI; 1987-362707/51.
XX New phosphopeptides contg. defined amino acid sequence - useful in
XX treatment of dental, rarefying bone diseases and disease relating to
XX malabsorption of minerals.
XX Claim 6; Page 17; 22pp; English.
XX The phosphopeptide is used in compsns. at a conc. of 0.01-5 wt%.
CC See also AAP1320-P1324.
XX Sequence 20 AA;
SQ Query Match 90.1%; Score 82; DB 8; Length 20;
Best Local Similarity 80.0%; Pred. No. 2.9e-08;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 NTMEHVXXXEESIIXQETK 21
DB 1 NTMEHVSSSESIISQETK 20
RESULT 9
AAR14449
ID AAR14449 standard; Protein; 20 AA.
XX AAR14449;
XX 10-JAN-1992 (first entry)
XX Phosphopeptide #3 for increasing mineral absorption.
XX caries; gingivitis; periodontal disease; osteoporosis; osteomalacia;
XX anaemia.
XX Synthetic.
XX Key Location/Qualifiers
XX Modified-site 7 /label= OTHER
XX /note= "phosphoserine (Pse)"
XX Modified-site 8 /label= OTHER
XX /note= "Pse"
XX Modified-site 9 /label= OTHER
XX /note= "Pse"
XX Modified-site 15 /label= OTHER
XX /note= "Pse"
XX US5015628-A.
XX 14-MAY-1991.
XX 03-AUG-1990; 90US-0563798.
XX 03-AUG-1990; 90US-0563798.
XX (UYME-) UNIV OF MELBOURNE.
XX (VICT-) VICTORIAN DAIRY INDUSTRY.
XX Reynolds EC;
XX WPI; 1991-316875/43.
XX Novel phosphopeptides - useful for treating dental diseases,
XX rarefying diseases or diseases relating to malabsorption of minerals
XX Claim 1; Column 11; 8pp; English.
XX This is an example of a highly generic formula for a phosphopeptide
XX of length 5-13 amino acids. The peptides can be made synthetically
XX (e.g. chemical synthesis or genetic engineering) or they can be
XX extracted from cereals, nuts or vegetables or by fractionating a
XX digest of casein, alpha-s-casein, beta-casein or a salt of it.
XX Compositions comprising the peptide may take the form of foodstuff
XX or confectionery, dentifrices, mouthwashes and preparations for
XX topical application to teeth or gingival tissue. The peptides
XX significantly increase absorption of calcium, phosphate and iron in
XX the gut. See AAR14447-R14451.
XX Sequence 20 AA;
SQ Query Match 90.1%; Score 82; DB 12; Length 20;
Best Local Similarity 80.0%; Pred. No. 2.9e-08;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 NTMEHVXXXEESIIXQETK 21
DB 1 NTMEHVSSSESIISQETK 20
RESULT 10
AAR28431
ID AAR28431 standard; peptide; 20 AA.
XX AAR28431;
XX 19-MAR-1993 (first entry)
XX Anticariogenic phosphopeptide.
XX Casein; metal ion; dietetic; purification; growth medium; dietary
XX supplement; fertiliser.
XX Key Location/Qualifiers
XX Modified-site 7 /note= "post-translationally phosphorylated serine"
XX Modified-site 8 /note= "post-translationally phosphorylated serine"
XX Modified-site 9 /note= "post-translationally phosphorylated serine"
XX Modified-site 15 /note= "post-translationally phosphorylated serine"
XX W09218526-A.
XX 29-OCT-1992.
XX 16-APR-1992; 92WO-AU00175.
XX 19-APR-1991; 91AU-0005706.
XX (UYME) UNIV MELBOURNE.
XX (VICT-) VICTORIAN DAIRY IND AUTHORITY.
XX Reynolds EC;
XX WPI; 1992-382039/46.
XX Selected phospho-peptide(s) prodn. having anticariogenic
XX activities etc. - comprises digesting soluble monovalent cation
XX salt of casein in soln. introducing di- or trivalent metal ion
XX and filtering through filter having mol. wt. exclusion limit
XX Claim 11; Page 13; 18pp; English.
XX The peptide may be prepd. by completely digesting casein in soln.
XX with a proteolytic enzyme, adding mineral acid to the soln. to
XX adjust the pH to 4.7, removing any precipitate, adding calcium
XX chloride to cause aggregation of the peptides in soln. and
XX separating the aggregated phosphopeptides. This method allows prodn.

CC of the phosphopeptide by industrial methods. The phosphopeptide
 CC has anticariogenic activity and may be used as a dietetic. The
 CC peptide may be used in a microbiological growth medium, as a dietary
 CC supplement or as a fertiliser.
 CC See also AAR28425-33.
 XX
 SQ Sequence 20 AA;

Query Match 90.1%; Score 82; DB 13; Length 20;
 Best Local Similarity 80.0%; Pred. No. 2.9e-08;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 NTMEHVXXXESIIIXQETK 21
 ID 1 NTMEHVSSSERSIIISQETK 20

RESULT 11
 AAR32933
 ID AAR32933 standard; Protein; 20 AA.
 XX
 AC AAR32933;

DT 02-JUL-1993 (first entry)
 DE Casein phosphopeptide #7.

KW Casein; phosphopeptide; dental calculus; salts; alkaline metal;
 KW alkaline earth metal; Zn/phosphopeptide complex; aggregate;
 KW anti-carries; anti-gingivitis.
 XX
 OS Synthetic.

FH Key Location/Qualifiers
 FT Modified-site 7
 FT Modified-site 8 /label= Phosphoserine
 FT Modified-site 9 /label= Phosphoserine
 FT Modified-site 15 /label= Phosphoserine
 FT Modified-site 15 /label= Phosphoserine
 XX
 PN W09303707-A.

XX
 PD 04-MAR-1993.

PF 21-AUG-1992; 92WO-AU00441.

PR 22-AUG-1991; 91US-0748344.

XX (UYME) UNIV MELBOURNE.
 XX (VICT-) VICTORIAN DAIRY IND AUTHORITY.

XX Reynolds EC;

XX WPI; 1993-093685/11.

XX Controlling dental calculus by treating teeth with oral compsns.
 XX - which contains phospho-peptide(s) having 5-40 amino-acyl
 XX residues

PS Claim 3; Page 19; 23pp; English.

XX The sequences given in AAR32927-35 are casein phosphopeptides which can
 CC be used to inhibit dental calculus. These peptides are pref. in the
 CC form of salts selected from alkaline metal, alkaline earth metal salts
 CC such as Na, Ca, Zn, Cu, Al, K, Sr, Mg and Ni salts. These peptides
 CC are pref. present as a Zn/phosphopeptide complex or aggregate. These
 CC peptides have anti-calculus potential, and are anti-carries and anti-
 CC gingivitis agents.

XX Sequence 20 AA;

Query Match 90.1%; Score 82; DB 14; Length 20;
 Best Local Similarity 80.0%; Pred. No. 2.9e-08;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 NTMEHVXXXESIIIXQETK 21
 ID 1 NTMEHVSSSERSIIISQETK 20

RESULT 12
 AAR32310
 ID AAR32310 standard; peptide; 20 AA.
 XX
 AC AAR32310;

DT 10-JUN-1993 (first entry)

DE Phosphopeptide derived from casein.

KW Dental; teeth; tartar control; brushite; calcium phosphate;
 KW hydroxyapatite; mouthwash; toothpaste.

XX Synthetic.

FH Key Location/Qualifiers
 FT Modified-site 7 /note= "post-translationally phosphorylated"
 FT Modified-site 8 /note= "post-translationally phosphorylated"
 FT Modified-site 9 /note= "post-translationally phosphorylated"
 FT Modified-site 15 /note= "post-translationally phosphorylated"
 FT Modified-site 15 /note= "post-translationally phosphorylated"

XX EP528458-A.

XX 24-FEB-1993.

XX 03-JUL-1992; 92EP-0202024.

XX 09-AUG-1991; 91GB-0017315.

XX (UNIL) UNILEVER PLC.
 XX (UNIL) UNILEVER NV.

XX Burger AR, Schick LA;

XX WPI; 1993-060322/08.

XX Phospho-peptide(s) for dental tartar control - are included in
 XX compsns. with pyrophosphate(s) or zinc salts to provide good
 XX inhibition

XX Disclosure; Page 13; 17pp; English.

XX The peptide sequence is that of a phosphopeptide prep. from a
 CC tryptic digest of casein. The peptide may be used with an anti-
 CC calculus agent, e.g. alkali metal pyrophosphate, in a compsn. for
 CC controlling dental tartar. The compsn. inhibits conversion of
 CC brushite and amorphous calcium phosphate into the more stable
 CC hydroxyapatite on the teeth. The compsn. is used in the form of a
 CC mouthwash, toothpaste, gel, lozenge or chewing gum, for care of the
 CC teeth and tartar control.
 CC See also AAR32308-12.

XX Sequence 20 AA;

Query Match 90.1%; Score 82; DB 14; Length 20;
 Best Local Similarity 80.0%; Pred. No. 2.9e-08;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 NTMEHVXXXESIIIXQETK 21

Db 1 NTMEHVSSSESIISQETK 20
 ||||| ||||| |||||

RESULT 13

AAR31239
 ID AAR31239 standard; Protein; 20 AA.

XX AC AAR31239;

XX DT 18-MAY-1993 (first entry)

XX DE Phosphopeptide 3.

XX KW Phosphopeptide; active agent; oral; composition: anionic; polymeric;
 KW stabiliser; carboxylate; polymer; sulfonate; destabilisation;
 KW fluoride; caries; gingivitis.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 7

FT /label= Phosphoserine

FT Modified-site 8

FT /label= Phosphoserine

FT Modified-site 9

FT /label= Phosphoserine

FT Modified-site 15

FT /label= Phosphoserine

FT FT

XX EP523776-A.

XX PN

XX PD 20-JAN-1993.

XX PF 03-JUL-1992; 92EP-0202023.

XX PS 17-JUL-1991; 91US-0731592.

XX PR (UNIL) UNILEVER NV.

XX PA (UNIL) UNILEVER PLC.

XX PI Burger AR, Elliott DL, Schick LA;

XX XX WPI; 1993-019802/03.

XX DR

XX XX

XX PT

XX PT

XX PT

XX PT

XX PS

XX XX

XX XX

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX XX

XX SQ

Query Match 90.1%; Score 82; DB 14; Length 20;

Best Local Similarity 80.0%; Pred. No. 2.9e-08;

Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTMEHVXXXEEESIISQETK 21

Db 1 NTMEHVSSSESIISQETK 20

||||| ||||| |||||

RESULT 14

AAR47820
 ID AAR47820 standard; peptide; 20 AA.

XX AC AAR47820;

XX DT 21-JUL-1994 (first entry)

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

FT Modified-site 7 /label= OTHER
 FT /note= "Phosphoserine"
 FT Modified-site 8
 FT /label= OTHER
 FT /note= "Phosphoserine"
 FT Modified-site 9
 FT /label= OTHER
 FT /note= "Phosphoserine"
 FT Modified-site 15
 FT /label= OTHER
 FT /note= "Phosphoserine"
 PN EP629393-A.
 XX
 XX 21-DEC-1994.
 XX
 PF 07-JUN-1994; 94EP-0304083.
 XX
 PR 16-JUN-1993; 93US-0078706.
 XX
 PA (ICIL) ICI AMERICAS INC.
 XX
 PI Duross JW;
 XX
 DR WPI; 1995-044845/07.
 XX
 XX Crystalline poly:ol compositions for use in confectionery and
 PT pharmaceuticals - contain a phospho:peptide uniformly
 PT distributed within the crystal matrix and can improve oral
 PT hygiene
 XX
 PS Disclosure; column 4; 8pp; English.
 XX
 CC Peptides (AAR68936-40) are phosphopeptides resulting from the tryptic
 CC digestion of sodium caseinate. The peptide shown here is derived from
 CC the TCPK-tryptic digest of alpha(s2)-caseinate (comprising alpha(s2),
 CC alpha(s3), alpha(s4) and alpha(s6)). The peptides contain the amino
 CC acids phosphoserine, phosphothreonine, phosphotyrosine or
 CC phosphohistidine, and include in their sequence aspartate and
 CC glutamate.
 CC The phosphopeptides form part of a novel composition containing the
 CC peptide dispersed uniformly in a crystal mix of sugars chosen from
 CC sorbitol, mannitol, xylitol, lactitol, cellobitol or mixtures of
 CC sorbitol/mannitol or sorbitol/xylitol. The compositions can be used
 CC instead of conventional sugars to manufacture confectionary products.
 CC The phosphopeptides are known to inhibit caries and gingivitis and can
 CC act as a source of calcium ions to promote/enhance remineralisation.
 XX
 SQ Sequence 20 AA;
 Query Match 90.1%; Score 82; DB 16; Length 20;
 Best Local Similarity 80.0%; Pred. No. 2.9e-08;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 NTMEHVXXXEEESIIXQETVK 21
 Db 1 NTMEHVSSESSESIISQETVK 20

Search completed: February 11, 2003, 18:16:56
 Job time : 31.5258 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:14:31 ; Search time 10.3918 Seconds
(without alignments)
59.459 Million cell updates/sec

Title: US-09-380-738A-4
Perfect score: 91
Sequence: 1 KNTMEHVXXXEESIIXQETVK 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	95.6	21	1	US-07-748-344B-8
2	87	95.6	21	2	US-08-954-985A-8
3	87	95.6	21	4	US-08-137-086-8
4	87	95.6	222	3	US-09-066-408-6
5	82	90.1	20	1	US-07-748-344B-7
6	82	90.1	20	2	US-08-954-985A-7
7	82	90.1	20	4	US-08-137-086-7
8	59.5	65.4	235	3	US-09-066-408-12
9	54.5	59.9	223	3	US-09-066-408-11
10	53.5	58.8	223	3	US-09-066-408-7
11	53.5	58.8	223	3	US-09-066-408-8
12	53.5	58.8	223	3	US-09-066-408-9
13	42	46.2	566	1	US-08-415-823-4
14	42	46.2	566	2	US-09-086-662-4
15	38	41.8	436	4	US-09-734-673-2
16	38	41.8	436	4	US-09-523-849-2
17	37	40.7	92	1	US-08-414-926A-17
18	37	40.7	92	2	US-08-926-922-17
19	37	40.7	92	3	US-09-253-682-17
20	37	40.7	92	4	US-09-527-657-17
21	36	39.6	285	4	US-09-134-001C-4756
22	36	39.6	315	4	US-09-134-001C-4771
23	36	39.6	2210	4	US-09-309-572-7
24	35	38.5	640	3	US-09-026-343-2
25	35	38.5	640	4	US-09-362-871-2
26	34	37.4	441	1	US-08-403-866-10
27	34	37.4	754	3	US-09-005-180A-3

28	33.5	36.8	534	2	US-08-317-401E-2
29	33.5	36.8	552	4	US-08-317-401E-4
30	33	36.3	198	4	US-09-134-001C-4395
31	33	36.3	215	2	US-08-852-809-4
32	33	36.3	548	6	5256558-5
33	33	36.3	796	3	US-09-005-180A-1
34	33	36.3	988	2	US-08-286-819A-19
35	33	36.3	988	3	US-08-980-357-19
36	32.5	35.7	108	3	US-08-966-269-15
37	32.5	35.7	108	4	US-09-436-183A-15
38	32	35.2	64	4	US-08-858-207A-517
39	32	35.2	340	2	US-08-446-875-16
40	32	35.2	449	3	US-08-462-351-2
41	32	35.2	449	4	US-09-602-807-2
42	32	35.2	449	6	5194425-1
43	32	35.2	624	2	US-08-879-561-9
44	31	34.1	181	4	US-08-858-207A-317
45	31	34.1	210	1	US-08-078-090-2

ALIGNMENTS

RESULT 1
US-07-748-344B-8
; Sequence 8, Application US/07748344B
; Patent No. 5227154
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; TITLE OF INVENTION: PHOSPHOPEPTIDES FOR THE
; TREATMENT OF DENTAL CALCULUS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDRUS, SCEALES, STARKE & SAWALL
; STREET: 100 EAST WISCONSIN AVE., SUITE 1100
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07748,344B
; FILING DATE: 19910822
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: SARA, CHARLES S
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: C.8493-87
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 255-2022
; TELEFAX: (608) 255-2182
; TELEX: 26832 ANDSTARK
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; FEATURE:
; FEATURE: 8 Phosphoserine
; FEATURE: 8 Post-translationally phosphorylated serine
; FEATURE: 8 Phosphoserine

Query Match 95.6%; Score 87; DB 3; Length 222;

US-0/-148-344B-1

Query Match 90.1%; Score 82; DB 1; Length 20;
Best Local Similarity 80.0%; Pred. No. 2.9e-09;
Matches 16; Conservative 0; Mismatches 4; Indels 0;

QY 2 NTMEHVXXXESIIIXQETK 21
Db 1 NTMEHVSSSESIIISQETK 20

RESULT 6
US-08-954-985A-7
; Sequence 7, Application US/08954985A
; Patent No. 5981475
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, Eric Charles
; TITLE OF INVENTION: A Treatment for Sensitive Teeth
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dental School, The University of Melbourne
; STREET: 711 Elizabeth Street
; CITY: Melbourne
; STATE: Victoria
; COUNTRY: Australia
; ZIP: 3000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Ascii
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,985A
; FILING DATE: 21-OCT-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,479
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: WOZNY, Thomas M
; REGISTRATION NUMBER: 28,922
; REFERENCE/DOCKET NUMBER: 322-00033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 271-7590
; TELEFAX: (414) 271-5770
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: Amino Acid
; TOPOLOGY: Linear
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 7
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 8
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 9
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 15
; OTHER INFORMATION: Post-translationally phosphorylated serine
; US-08-954-985A-7

Query Match 90.1%; Score 82; DB 2; Length 20;
Best Local Similarity 80.0%; Pred. No. 2.9e-09;
Matches 16; Conservative 0; Mismatches 4; Indels 0;

QY 2 NTMEHVXXXESIIIXQETK 21
Db 1 NTMEHVSSSESIIISQETK 20

RESULT 7
US-08-137-086-7
; Sequence 7, Application US/08137086
; Patent No. 6448374
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; TITLE OF INVENTION: PRODUCTION OF PHOSPHOPEPTIDES FROM CASEIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDRUS, SCEALES, STARKE & SAWALL
; STREET: 100 EAST WISCONSIN AVE., SUITE 1100
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,086
; FILING DATE: 04-MAR-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: SARA, CHARLES S
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: C.8493-87
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 255-2022
; TELEFAX: (608) 255-2182
; TELEX: 26832 ANDSTARK
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: Amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; FEATURE:
; FEATURE: Phosphoserine
; FEATURE: 7
; FEATURE:
; FEATURE: Post-translationally phosphorylated serine
; FEATURE:
; FEATURE: Phosphoserine
; FEATURE: 8
; FEATURE:
; FEATURE: Post-translationally phosphorylated serine
; FEATURE:
; FEATURE: Phosphoserine
; FEATURE: 9
; FEATURE:
; FEATURE: Post-translationally phosphorylated serine
; FEATURE:
; FEATURE: Phosphoserine
; FEATURE: 15
; FEATURE:
; FEATURE: Post-translationally phosphorylated serine
; US-08-137-086-7

Query Match 90.1%; Score 82; DB 4; Length 20;
Best Local Similarity 80.0%; Pred. No. 2.9e-09;
Matches 16; Conservative 0; Mismatches 4; Indels 0;

QY 2 NTMEHVXXXESIIIXQETK 21
Db 1 NTMEHVSSSESIIISQETK 20

```

; APPLICANT: Liu, Qing-Ming
; TITLE OF INVENTION: Casein Fragments Having Growth Promoting
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,408
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/02658
; FILING DATE: 31-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9522302.0
; FILING DATE: 31-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 018317-000100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..223
; OTHER INFORMATION: /note= "sheep alpha-S2 casein precursor"
;
; US-09-066-408-11
;
; Query Match          59.9%   Score 54.5; DB 3; Length 223;
; Best Local Similarity 59.1%; Pred. No. 0.0061;
; Matches 13; Conservative 1; Mismatches 7; Indels 1; Gaps 1;
;
; Qy      1 KNTMEHVXXXEESI-IXQETKY 21
;         | : |||| ||| | | | | |
; Db     16 KHKMEHVSSEPINISQEIVK 37
;
; RESULT 10
; US-09-066-408-7
; ; Sequence 7, Application US/09066408
; ; Patent No. 6060448
; ; GENERAL INFORMATION:
; ; - APPLICANT: Smith, John Arthur
; ; APPLICANT: Wilkinson, Mark Charles
; ; APPLICANT: Liu, Qing-Ming
; ; TITLE OF INVENTION: Casein Fragments Having Growth Promoting
; ; NUMBER OF SEQUENCES: 12
; ; CORRESPONDENCE ADDRESS:
; ; ADDRESSEE: Townsend and Townsend and Crew LLP
; ; STREET: Two Embarcadero Center, Eighth Floor
; ; CITY: San Francisco
; ; STATE: California
; ; COUNTRY: USA

```

```

; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,408
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/02658
; FILING DATE: 31-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9522302.0
; FILING DATE: 31-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 018317-0001000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..223
; OTHER INFORMATION: /note= "goat alpha-s2 casein
; precursor allele A"
;
US-09-066-408-7

```

```

Query Match 58.8%; Score 53.5; DB 3; Length 223;
Best Local Similarity 59.1%; Pred. No. 0.0094;
Matches 13; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

QY 1 KNTMEHVXXXXESI-IXQETVK 21
|: |||| | | | | |
Db 16 KHKMEHVSSEEPINIFQEIYK 37

RESULT 11
US-09-066-408-8
; Sequence 8, Application US/09066408
; Patent No. 6060448
; GENERAL INFORMATION:
; APPLICANT: Smith, John Arthur
; APPLICANT: Wilkinson, Mark Charles
; APPLICANT: Liu, Qing-Ming
; TITLE OF INVENTION: Casein Fragments Having Growth Promoting
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,408
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 514
;

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/02658
; FILING DATE: 31-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9522302.0
; FILING DATE: 31-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 018317-0001000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..223
; OTHER INFORMATION: /note= "goat alpha-s2 casein
; precursor allele B"
;
US-09-066-408-8

```

```

Query Match 58.8%; Score 53.5; DB 3; Length 223;
Best Local Similarity 59.1%; Pred. No. 0.0094;
Matches 13; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

QY 1 KNTMEHVXXXXESI-IXQETVK 21
|: |||| | | | | |
Db 16 KHKMEHVSSEEPINIFQEIYK 37

```

```

RESULT 12
US-09-066-408-9
; Sequence 9, Application US/09066408
; Patent No. 6060448
; GENERAL INFORMATION:
; APPLICANT: Smith, John Arthur
; APPLICANT: Wilkinson, Mark Charles
; APPLICANT: Liu, Qing-Ming
; TITLE OF INVENTION: Casein Fragments Having Growth Promoting
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/066,408
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/02658
; FILING DATE: 31-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9522302.0
; FILING DATE: 31-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 018317-0001000US
;

```


TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..223
OTHER INFORMATION: /note= "goat alpha-s2 casein precursor allele C"
US-09-066-408-9
Query Match 58.8%; Score 53.5; DB 3; Length 223;
Best Local Similarity 59.1%; Pred. No. 0.0094; 7; Indels 1; Gaps 1;
Matches 13; Conservative 1; Mismatches 7;
QY 1 KNTMEHVXXXEESI-IXQETVK 21
Db 16 KHKMEHVSSSEEPINIFQETVK 37
RESULT 13
US-08-415-823-4
Sequence 4, Application US/08415823
Patent No. 5759538
GENERAL INFORMATION:
APPLICANT: Donovan, William P.
TITLE OF INVENTION: BACILLUS THURINGIENSIS apr AND npr
TITLE OF INVENTION: PROTEASE GENES AND ALKALINE PROTEASE DEFICIENT AND NEUTRAL
TITLE OF INVENTION: PROTEASE DEFICIENT B.T. STRAINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Panitch Schwarze Jacobs & Nadel, P.C. c/o
STREET: 1601 Market Street, 36th floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103-2398
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/415,823
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/415,823
FILING DATE:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 566 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-086-662-4
Query Match 46.2%; Score 42; DB 2; Length 566;
Best Local Similarity 40.0%; Pred. No. 3.8;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 2 NTMEHVXXXEESIIXQETVK 21
Db 221 NTIDHTNDKSPVKQEPK 240
RESULT 15
US-09-734-673-2
Sequence 2, Application US/09734673
Patent No. 6410294
GENERAL INFORMATION:
APPLICANT: GUEGLER, Karl et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001020
CURRENT APPLICATION NUMBER: US/09/734,673
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 436
TYPE: PRT
ORGANISM: Human
US-09-734-673-2
Query Match 41.8%; Score 38; DB 4; Length 436;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 KNTMEHVXXXEESI 14

Wed Feb 12 11:35:24 2003

us-09-380-738a-4.rai

Page 8

Db 358 KSTLRLGITEHI 371
|:|:|: | | |

Search completed: February 11, 2003, 18:22:10
Job time : 11.5918 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:19:51 ; Search time 10.1753 Seconds
(without alignments)
52.729 Million cell updates/sec

Title: US-09-380-738A-4

Perfect score: 91

Sequence: 1 KNTMEHVXXXEESIIXQETK 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	37	40.7	121	US-09-741-669-447	Sequence 447, App
2	37	40.7	879	US-10-108-605-217	Sequence 217, App
3	36	39.6	666	US-09-815-242-13637	Sequence 13637, A
4	36	39.6	676	US-09-815-242-13513	Sequence 13513, A
5	35	38.5	54	US-09-864-761-47515	Sequence 47515, A
6	35	38.5	415	US-09-815-242-12593	Sequence 12593, A
7	35	38.5	415	US-09-815-242-12593	Sequence 12593, A
8	35	38.5	487	US-09-815-242-5509	Sequence 5509, App
9	35	38.5	640	US-10-028-780-2	Sequence 2, Appli
10	35	38.5	677	US-09-925-300-1626	Sequence 1626, App
11	34	37.4	1116	US-09-790-318-2	Sequence 2, Appli
12	33.5	36.8	345	US-09-738-626-3894	Sequence 3894, App
13	33	36.3	565	US-09-925-301-1302	Sequence 1302, App
14	33	36.3	693	US-09-789-919-48	Sequence 48, Appl
15	33	36.3	813	US-09-764-898-197	Sequence 197, App
16	32	35.2	154	US-09-738-626-5286	Sequence 5286, App
17	32	35.2	394	US-09-815-242-13500	Sequence 13500, A
18	32	35.2	599	US-09-952-013A-4	Sequence 4, Appli
19	32	35.2	740	US-09-815-242-10876	Sequence 10876, A

```

20 32 35.2 1029 10 US-09-815-242-5885 Sequence 5885, App
21 32 35.2 1048 10 US-09-815-242-13083 Sequence 13083, A
22 32 35.2 2086 10 US-09-815-242-5639 Sequence 5639, App
23 32 35.2 5795 10 US-09-815-242-12610 Sequence 12610, A
24 32 35.2 6281 10 US-09-815-242-12996 Sequence 12996, A
25 31 34.1 164 10 US-09-815-242-11411 Sequence 11411, A
26 31 34.1 164 10 US-09-815-242-11573 Sequence 11573, A
27 31 34.1 169 10 US-09-925-301-1684 Sequence 1684, App
28 31 34.1 172 9 US-09-922-199A-26 Sequence 26, Appli
29 31 34.1 353 9 US-09-971-228-9 Sequence 9, Appli
30 31 34.1 353 10 US-09-771-063-2 Sequence 2, Appli
31 31 34.1 353 10 US-09-771-063-4 Sequence 4, Appli
32 31 34.1 353 10 US-09-842-316-8 Sequence 8, Appli
33 31 34.1 353 10 US-09-731-030A-19 Sequence 19, Appli
34 31 34.1 353 12 US-10-037-616-23 Sequence 23, Appli
35 31 34.1 354 10 US-09-925-302-501 Sequence 501, App
36 31 34.1 391 10 US-09-934-249-7 Sequence 7, Appli
37 31 34.1 398 10 US-09-925-302-472 Sequence 472, App
38 31 34.1 411 10 US-09-815-242-13666 Sequence 13666, A
39 31 34.1 438 10 US-09-997-701-2 Sequence 2, Appli
40 31 34.1 440 9 US-10-036-041-9 Sequence 9, Appli
41 31 34.1 440 9 US-10-035-855-9 Sequence 9, Appli
42 31 34.1 440 9 US-10-174-590-442 Sequence 442, App
43 31 34.1 440 9 US-10-176-738-442 Sequence 442, App
44 31 34.1 440 9 US-10-175-737-442 Sequence 442, App
45 31 34.1 440 9 US-10-173-706-442 Sequence 442, App

```

ALIGNMENTS

```

RESULT 1
US-09-741-669-447
; Sequence 447, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Alllyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; proliferation of E. coli
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 447
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-447

```

```

Query Match 40.7%; Score 37; DB 10; Length 121;
Best Local Similarity 38.9%; Pred. No. 4.8;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 KNTMEHVXXXEESIIXQE 18
| | | | | | | | | | | | | | | |
DB 104 KLOMHVDPSEINIVVQK 121

```

```

RESULT 2
US-10-108-605-217
; Sequence 217, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Bachmann, Lynn
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT EN

```



```

; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47515
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL157791.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
; OTHER INFORMATION: EST HUMAN HIT: AL040793.1, EVALUATE 8.00e-24
; OTHER INFORMATION: SWISSPROT HIT: Q39565, EVALUATE 2.70e+00
US-09-864-761-47515

```

```

Query Match      38.5%; Score 35; DB 10; Length 54;
Best Local Similarity 46.2%; Pred. No. 4.7;
Matches 6; Conservative 3; Mismatches 4; Indels 4; Gaps 0;

```

```

QY 2 NTMEHVXXXEESI 14
||||| : : :
DB 41 NTMEHLSLLDNNI 53

```

```

RESULT 6
US-09-815-242-12593
; Sequence 12593, Application US/09/815,242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12593
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727

```

```

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12593
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12593

```

```

Query Match      38.5%; Score 35; DB 10; Length 415;
Best Local Similarity 33.3%; Pred. No. 43;
Matches 6; Conservative 4; Mismatches 8; Indels 8; Gaps 0;

```

```

QY 4 MEHVXXXEESIIXQETK 21
||||| : : :
DB 1 MEHVFYHRHDVSNESYQ 18

```

```

RESULT 7
US-09-815-242-12909
; Sequence 12909, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12909
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12909

```

```

Query Match      38.5%; Score 35; DB 10; Length 415;
Best Local Similarity 33.3%; Pred. No. 43;
Matches 6; Conservative 4; Mismatches 8; Indels 8; Gaps 0;

```

```

QY 4 MEHVXXXEESIIXQETK 21
||||| : : :
DB 1 MEHVFYHRHDVSNESYQ 18

```

Query Match 38.5%; Score 35; DB 10; Length 677;

Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 2 NTMEHVXXXEESIIXQETY 20
| | : | |
db 12 NRCEHLERTIESILNQ-TY 29

```

RESULT 13
US-09-925-301-1302
; Sequence 1302, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL06
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1302
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1302

```

```
Query Match          36.3%; Score 33; DB 10; Length 565;
Best Local Similarity 36.8%; Pred. No. 1.4e+02;
Matches 7; Conservative      3; Mismatches 9; Indels 0; Gaps 0;
```

Qy 2 NTMEHVXXXEESIIXQETY 20
 !: | : : | | |
 Db 320 NSEEGIOALDEVIFSLEY 338

```

RESULT 14
US-09-789-919-48
; Sequence 48, Application US/09789919
; Patent No. US20020064855A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor
; APPLICANT: Moore, Kateri
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; TITLE OF INVENTION: CELLS AND USES THEREOF
; FILE REFERENCE: 2275-1-005
; CURRENT APPLICATION NUMBER: US/09/789,919
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
; LENGTH: 693
; TYPE: prt
; ORGANISM: Mus musculus
US-09-789-919-48

```

Query Match	36.3%	Score 33;	DB 10;	Length 693;
Best Local Similarity	33.3%	Pred. No. 1.8e+02;		
Matches	7;	Conservative	4;	Mismatches 10;
			Indels	0;
			Gaps	0;

QY 1 KNTMEHVXXXEESIIXQETK 21
 | : | : | : | : |
Dh 355 KGAIGHIVPTSEKSILAVEKNK 375

RESULT 15
US-09-764-898-197
; Sequence 197, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:

Best Local Similarity	42.1%;	Pred. No. 74;	
Matches	8;	Conservative	3;
Mismatches	8;	Indels	8;
Gaps	0;		

09Y 3 TMEHVXXXEESIIXQETK 21
I : I I I : I I I :
db 65 TVLHVKITETAIRALETYO 83

```

RESULT 11
US-09-790-318-2
; sequence 2, Application US/09790318
; Patent No. US20020038012A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; TITLE OF INVENTION: No. US20020038012A1e1 Human Regulatory Protein and Polynucleotide
; FILE REFERENCE: LEX-0136-USA
; CURRENT APPLICATION NUMBER: US/09/790,318
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 60/184,015
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1116
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-790-318-2

```

Query Match 37.4%; Score 34; DB 10; Length 1116;
Best Local Similarity 36.8%; Pred. No. 2e+02;
Matches 7: Conservative 4: Mismatches 8: Indels 0: Gaps 0;

QY 1 KNTMEHVXXXEESIIXQET 19
:| | | | :
pb 1055 ENTKNSVDFOEFTVISES 1073

```

RESULT 12
US-09-738-626-3894
; Sequence 3894, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3894
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3894

```

Query Match 36.88; Score 33.5; DB 9; Length 345;

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P201
; CURRENT APPLICATION NUMBER: US/09/764,898
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 197
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-898-197

Query Match 36.3%; Score 33; DB 10; Length 813;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTMEHVXXEES 13
||:|:
Db 789 NTLEHLRLRES 800

Search completed: February 11, 2003, 18:36:16
Job time : 11.1753 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:17:01 ; Search time 144.186 Seconds
(without alignments)
93.903 Million cell updates/sec

Title: US-09-380-738A-4

Perfect score: 91

Sequence: 1 KNTMEHVXXXXEIIIXOETVK 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pcp.*
- 2: /cgn2_6/ptodata/1/paa/US07_COMB.pcp.*
- 3: /cgn2_6/ptodata/1/paa/US06_COMB.pcp.*
- 4: /cgn2_6/ptodata/1/paa/US080_COMB.pcp.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pcp.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pcp.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pcp.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pcp.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pcp.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pcp.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pcp.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pcp.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pcp.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pcp.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pcp.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pcp.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pcp.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pcp.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pcp.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pcp.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pcp.*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pcp.*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pcp.*
- 24: /cgn2_6/ptodata/1/paa/US100_COMB.pcp.*
- 25: /cgn2_6/ptodata/1/paa/US101_COMB.pcp.*
- 26: /cgn2_6/ptodata/1/paa/US102_COMB.pcp.*
- 27: /cgn2_6/ptodata/1/paa/US60_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Length	ID	Description
1	83	91.2	21	17	US-09-380-738A-4
2	82	90.1	20	3	US-07-731-592B-3
3	45	49.5	279	19	US-09-573-655A-1509
4	45	49.5	279	19	US-09-573-655B-1509
5	43	47.3	408	1	PCT-US01-08631-47481
6	43	47.3	921	1	PCT-US01-42950-533

7	43	47.3	942	1	PCT-US01-42950-532
8	42	46.2	566	21	US-09-791-537-91913
9	42	46.2	566	21	US-09-791-537-125976
10	41	45.1	187	17	US-09-328-352-5117
11	40	44.0	861	19	US-09-540-209B-8326
12	39.5	43.4	295	15	US-09-198-452A-1123
13	39.5	43.4	339	18	US-09-438-185-1049
14	39.5	43.4	339	18	US-09-438-185A-1049
15	39	42.9	409	22	US-09-897-516-6405
16	39	42.9	409	22	US-09-897-516-6405
17	39	42.9	1458	27	US-60-360-039-2348
18	38	41.8	2206	1	PCT-US99-26796-111
19	38	41.8	130	27	US-60-173-468-1355
20	38	41.8	132	1	PCT-US01-01334-3437
21	38	41.8	132	21	US-09-764-874-3437
22	38	41.8	183	27	US-60-182-567-689
23	38	41.8	196	27	US-60-169-868-5211
24	38	41.8	242	20	US-09-614-150-2715
25	38	41.8	242	27	US-60-167-217-2763
26	38	41.8	242	27	US-60-173-464-2243
27	38	41.8	242	27	US-60-191-637-2721
28	38	41.8	242	27	US-60-191-681-2172
29	38	41.8	382	27	US-60-143-188-1
30	38	41.8	435	1	PCT-US02-14460-2
31	38	41.8	435	25	US-10-142-356-2
32	38	41.8	435	27	US-60-290-276-2
33	38	41.8	436	1	PCT-US00-07715-2
34	38	41.8	436	25	US-10-195-101-2
35	38	41.8	891	18	US-09-451-320-2418
36	38	41.8	891	21	US-09-791-537-121930
37	38	41.8	891	27	US-60-389-987-565
38	38	41.8	891	27	US-60-412-418-565
39	38	41.8	1846	1	PCT-US02-22866-29
40	37	40.7	55	18	US-09-450-969-4451
41	37	40.7	92	12	US-08-812-716-14
42	37	40.7	92	22	US-09-892-100-17
43	37	40.7	121	21	US-09-741-669-447
44	37	40.7	132	18	US-09-489-039A-12298
45	37	40.7	149	19	US-09-513-996A-6781

ALIGNMENTS

RESULT 1
US-09-380-738A-4
; Sequence 4, Application US/09380738A
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, Eric
; TITLE OF INVENTION: CALCIUM PHOSPHOPEPTIDE COMPLEXES
; FILE REFERENCE: 040268/0161
; CURRENT APPLICATION NUMBER: US/09/380,738A
; CURRENT FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: PCT/AU98/00160
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: AU P05662
; PRIOR FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Bos sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: Xaa is a phosphorylated Serine
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: Xaa is a phosphorylated Serine
; NAME/KEY: misc_feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: Xaa is a phosphorylated Serine

; NAME/KEY: misc_feature
; LOCATION: (16)..(16)
; OTHER INFORMATION: Xaa is a phosphorylated Serine
US-09-380-738A-4

Query Match 91.2%; Score 83; DB 17; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXEESIIXQETK 21
|||||
Db 1 KNTMEHVXXXEESIIXQETK 21

RESULT 2

US-07-731-592B-3

; Sequence 3, Application US/07731592B
; GENERAL INFORMATION:
; APPLICANT: Burger, Allan R.
; APPLICANT: Elliott, David L.
; APPLICANT: Schick, Laura A.
; TITLE OF INVENTION: Oral Compositions Containing a
; TITLE OF INVENTION: Phosphopeptide
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Department, Unilever United States, Inc.
; STREET: 45 River Road
; CITY: Edgewater
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette-3.5 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word(R)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/731,592B
; FILING DATE: 19910717
; CLASSIFICATION: 514
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: AMINO ACID
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 7
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 8
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 9
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 15
; OTHER INFORMATION: Post-translationally phosphorylated
; OTHER INFORMATION: serine
US-07-731-592B-3

Query Match 90.1%; Score 82; DB 3; Length 20;
Best Local Similarity 80.0%; Pred. No. 1e-08;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTMEHVXXXEESIIXQETK 21

||||| ||||| ||||| |||||
Db 1 NTMEHVSSSEESIISQETK 20

RESULT 3

US-09-573-655A-1509
; Sequence 1509, Application US/09573655A
; GENERAL INFORMATION:
; APPLICANT: SOLOVYEV, Victor and TROUKHAN, Maxim
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypept
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-0876P
; CURRENT APPLICATION NUMBER: US/09/573,655A
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 3280
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1509
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-573-655A-1509

Query Match 49.5%; Score 45; DB 19; Length 279;
Best Local Similarity 41.2%; Pred. No. 6.4;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXEESIIXQ 17
||||| ||||| ||||| |||||
Db 128 KATMDHIAAAEEQIVSE 144

RESULT 4

US-09-573-655B-1509
; Sequence 1509, Application US/09573655B
; GENERAL INFORMATION:
; APPLICANT: SOLOVYEV, Victor and TROUKHAN, Maxim
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypept
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-0876P
; CURRENT APPLICATION NUMBER: US/09/573,655B
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 3281
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1509
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-573-655B-1509

Query Match 49.5%; Score 45; DB 19; Length 279;
Best Local Similarity 41.2%; Pred. No. 6.4;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXEESIIXQ 17
||||| ||||| ||||| |||||
Db 128 KATMDHIAAAEEQIVSE 144

RESULT 5

PCT-US01-08631-47481
; Sequence 47481, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom

; SEQ ID NO 47481
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-08631-47481

Query Match 47.3%; Score 43; DB 1; Length 408;
Best Local Similarity 42.1%; Pred. No. 26;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 TMEHVXXXEESIIXQETK 21
|: || | | | | | | | | |
Db 253 TVNHRFSENEIIIEDDYK 271

RESULT 6

PCT-US01-42950-533
; Sequence 533, Application PC/TUS0142950

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-096
; CURRENT APPLICATION NUMBER: PCT/US01/42950
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/714,936
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 533
; LENGTH: 921
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-42950-533

Query Match 47.3%; Score 43; DB 1; Length 921;
Best Local Similarity 42.1%; Pred. No. 73;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 TMEHVXXXEESIIXQETK 21
|: || | | | | | | | | |
Db 84 TVNHRFSENEIIIEDDYK 102

RESULT 7

PCT-US01-42950-532
; Sequence 532, Application PC/TUS0142950

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-096
; CURRENT APPLICATION NUMBER: PCT/US01/42950
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/714,936
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 532
; LENGTH: 942
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-42950-532

Query Match 47.3%; Score 43; DB 1; Length 942;
Best Local Similarity 42.1%; Pred. No. 76;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 TMEHVXXXEESIIXQETK 21
|: || | | | | | | | | |
Db 84 TVNHRFSENEIIIEDDYK 102

RESULT 8

US-09-791-537-91913

; Sequence 91913, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.

; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 91913
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-09-791-537-91913

Query Match 46.2%; Score 42; DB 21; Length 566;
Best Local Similarity 40.0%; Pred. No. 63;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 NTMEHVXXXEESIIXQETK 21
|: || | | | | | | | | |
Db 221 NTIDHVTNDDKSPVKQEAPK 240

RESULT 9

US-09-791-537-125976
; Sequence 125976, Application US/09791537

; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 125976
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Bacillus cereus
US-09-791-537-125976

Query Match 46.2%; Score 42; DB 21; Length 566;
Best Local Similarity 40.0%; Pred. No. 63;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 NTMEHVXXXEESIIXQETK 21
|: || | | | | | | | | |
Db 221 NTIDHVTNDDKSPVKQEAPK 240

RESULT 10

US-09-328-352-5117
; Sequence 5117, Application US/09328352

; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5117
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5117

```

Query Match          45.1%; Score 41; DB 17; Length 187;
Best Local Similarity 42.9%; Pred. No. 24;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXEESIIXQETK 21
   ||: | ||||: :| |
Db 101 KNSEEDFKTKESILLNDTLK 121

RESULT 11
US-09-540-209B-8326
; Sequence 8326, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 8326
; LENGTH: 861
; TYPE: PRT
; ORGANISM: B.fragilis
US-09-540-209B-8326

Query Match          44.0%; Score 40; DB 19; Length 861;
Best Local Similarity 42.1%; Pred. No. 2.7e+02;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXEESIIXQET 19
   | :|| :|||: |||
Db 347 KGD LKHGNYQESILNQET 365

RESULT 12
US-09-198-452A-1123
; Sequence 1123, Application US/09198452A
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 1123
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-1123

Query Match          43.4%; Score 39.5; DB 15; Length 295;
Best Local Similarity 45.0%; Pred. No. 86;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 2 NTMEHVXXXEESIIXQETK 21
   ||: | :|| | :|| |
Db 147 NTVPHIVGEEKIL-RETVK 165

RESULT 13
US-09-438-185-1049
; Sequence 1049, Application US/09438185
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia pneumoniae Genome Sequence
; FILE REFERENCE: 018941-00041105

```

```

; CURRENT APPLICATION NUMBER: US/09/438,185
; CURRENT FILING DATE: 1999-11-11
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1049
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-438-185-1049

```

```

Query Match          43.4%; Score 39.5; DB 18; Length 339;
Best Local Similarity 45.0%; Pred. No. 1e+02;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

```

```

QY 2 NTMEHVXXXEESIIXQETK 21
   ||: | :|| | :|| |
Db 191 NTVPHIVGEEKIL-RETVK 209

```

```

RESULT 14
US-09-438-185A-1049
; Sequence 1049, Application US/09438185A
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia pneumoniae Genome Sequence
; FILE REFERENCE: 018941-00041105
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1049
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn1048
US-09-438-185A-1049

```

```

Query Match          43.4%; Score 39.5; DB 18; Length 339;
Best Local Similarity 45.0%; Pred. No. 1e+02;
Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

```

```

QY 2 NTMEHVXXXEESIIXQETK 21
   ||: | :|| | :|| |
Db 191 NTVPHIVGEEKIL-RETVK 209

```

```

RESULT 15
US-09-897-516-6405
; Sequence 6405, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B

```

us-09-380-738a-4.rapm

Wed Feb 12 11:35:24 2003

```
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 6405
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516-6405
```

```
Query Match 42.9%; Score 39; DB 22; Length 409;
Best Local Similarity 33.3%; Pred. No. 1.6e+02;
Matches 7; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 1 KNTMEHVXXXEESIXQETK 21
| | | | : | : | |
Db 192 KYTMEKILTADRTIVRNDTQK 212
```

Search completed: February 11, 2003, 18:33:25
Job time : 146.186 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	40	44.0	2831	5	US-09-724-676-63963		Sequence 63963, A
2	40	44.0	2831	5	US-09-724-676A-63963		Sequence 63963, A
3	38	41.8	503	6	US-10-103-140-2		Sequence 2, Appli
4	38	41.8	1912	1	PCT-US02-21361-48		Sequence 48, Appl
5	38	41.8	1912	6	US-10-189-186-48		Sequence 48, Appl
6	37	40.7	803	5	US-09-724-676A-52616		Sequence 52616, A
7	37	40.7	803	5	US-09-724-676-52616		Sequence 52616, A
8	37	40.7	986	5	US-09-724-676-52629		Sequence 52629, A
9	37	40.7	986	5	US-09-724-676A-52629		Sequence 52629, A
10	37	40.7	1006	5	US-09-724-676-52622		Sequence 52622, A
11	37	40.7	1006	5	US-09-724-676-52635		Sequence 52635, A
12	37	40.7	1006	5	US-09-724-676-52639		Sequence 52639, A
13	37	40.7	1006	5	US-09-724-676A-52632		Sequence 52632, A
14	37	40.7	1006	5	US-09-724-676A-52635		Sequence 52635, A
15	37	40.7	1006	5	US-09-724-676A-52639		Sequence 52639, A
16	37	40.7	1007	5	US-09-724-676-52626		Sequence 52626, A
17	37	40.7	1007	5	US-09-724-676A-52626		Sequence 52626, A
18	37	40.7	1052	5	US-09-724-676-52632		Sequence 52632, A
19	37	40.7	1052	5	US-09-724-676A-52632		Sequence 52632, A
20	37	40.7	1088	5	US-09-724-676-52618		Sequence 52618, A
21	37	40.7	1088	5	US-09-724-676A-52618		Sequence 52618, A
22	36	39.6	153	5	US-09-724-676-70693		Sequence 70693, A
23	36	39.6	153	5	US-09-724-676A-70693		Sequence 70693, A
24	36	39.6	285	6	US-10-092-411A-4756		Sequence 4756, Ap
25	36	39.6	307	6	US-10-092-1410-5000		Sequence 5000, Ap
26	36	39.6	315	6	US-10-092-411A-4771		Sequence 4771, Ap

Db 1484 TLSEVKLEVETVIKQETY 1501

RESULT 3

US-10-103-140-2

; Sequence 2, Application US/10103140

; GENERAL INFORMATION:

; APPLICANT: LI, Li

; APPLICANT: LI, V. Yang

; TITLE OF INVENTION: Hemogen-EDAG: Novel Nuclear Factors Expressed in Hematopoietic

; TITLE OF INVENTION: Development

; FILE REFERENCE: 38368-179726

; CURRENT APPLICATION NUMBER: US/10/103,140

; CURRENT FILING DATE: 2002-12-30

; PRIOR APPLICATION NUMBER: US 60/277,624

; PRIOR FILING DATE: 2001-03-22

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 503

; TYPE: PRT

; ORGANISM: Mus musculus

US-10-103-140-2

Query Match 41.8%; Score 38; DB 6; Length 503;

Best Local Similarity 40.0%; Pred. No. 31;

Matches 8; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 2 NTMEHVXXXEESIIXQETK 21

:| | | | |

Db 352 STKTHQEAPEPEFISHETK 371

RESULT 4

PCT-US02-21361-48

; Sequence 48, Application PC/TUS0221361

; GENERAL INFORMATION:

; APPLICANT: Curagen Corporation, et al.

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 21402-397C

; CURRENT APPLICATION NUMBER: PCT/US02/21361

; CURRENT FILING DATE: 2002-07-03

; PRIOR APPLICATION NUMBER: 60/303046

; PRIOR FILING DATE: 2001-07-05

; PRIOR APPLICATION NUMBER: 60/360814

; PRIOR FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: 60/303828

; PRIOR FILING DATE: 2001-09-07

; PRIOR APPLICATION NUMBER: 60/323380

; PRIOR FILING DATE: 2001-09-19

; PRIOR APPLICATION NUMBER: 60/361133

; PRIOR FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: 60/304016

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: 60/304502

; PRIOR FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: 60/305262

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: 60/373881

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/305673

; PRIOR FILING DATE: 2001-07-16

Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 368

; SOFTWARE: Custom

; SEQ ID NO 48

; LENGTH: 1912

; TYPE: PRT

; ORGANISM: Homo sapiens

PCT-US02-21361-48

Query Match 41.8%; Score 38; DB 1; Length 1912;

Best Local Similarity 47.6%; Pred. No. 1.2e+02;

Matches 10; Conservative 2; Mismatches 3; Indels 6; Gaps 1;

QY 1 KNTMEHVXXXEESIIXQETK 21

||| ||||: |||:

Db 1463 KNT-----YEEISVQETLR 1477

RESULT 5

US-10-188-186-48

; Sequence 48, Application US/10188186

; GENERAL INFORMATION:

; APPLICANT: Anderson et al.

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 21402-397C

; CURRENT APPLICATION NUMBER: US/10/188,186

; CURRENT FILING DATE: 2002-07-02

; PRIOR APPLICATION NUMBER: 60/303046

; PRIOR FILING DATE: 2001-07-05

; PRIOR APPLICATION NUMBER: 60/360814

; PRIOR FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: 60/303828

; PRIOR FILING DATE: 2001-09-07

; PRIOR APPLICATION NUMBER: 60/323380

; PRIOR FILING DATE: 2001-09-19

; PRIOR APPLICATION NUMBER: 60/361133

; PRIOR FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: 60/304016

; PRIOR FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: 60/304502

; PRIOR FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: 60/305262

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: 60/373881

; PRIOR FILING DATE: 2002-04-19

; PRIOR APPLICATION NUMBER: 60/305673

; PRIOR FILING DATE: 2001-07-16

Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 368

; SOFTWARE: Custom

; SEQ ID NO 48

; LENGTH: 1912

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-188-186-48

Query Match 41.8%; Score 38; DB 6; Length 1912;

Best Local Similarity 47.6%; Pred. No. 1.2e+02;

Matches 10; Conservative 2; Mismatches 3; Indels 6; Gaps 1;

QY 1 KNTMEHVXXXEESIIXQETK 21

||| ||||: |||:

Db 1463 KNT-----YEEISVQETLR 1477

RESULT 6

US-09-724-676-52616

; Sequence 52616, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 52616

; LENGTH: 803

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676-52616

Query Match

Best Local Similarity 40.7%; Score 37; DB 5; Length 803;

Best Local Similarity 35.0%; Pred. No. 76;

Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXXEEIIIXQETY 20
:||||: : ||||
Db 330 RNTLEKFTASIQRLIEQEEY 349

RESULT 7

US-09-724-676A-52616
; Sequence 52616, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 52616

; LENGTH: 803

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676A-52616

Query Match 40.7%; Score 37; DB 5; Length 803;

Best Local Similarity 35.0%; Pred. No. 76;

Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXXEEIIIXQETY 20
:||||: : ||||
Db 330 RNTLEKFTASIQRLIEQEEY 349

RESULT 8

US-09-724-676-52629

; Sequence 52629, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 52629

; LENGTH: 986

; TYPE: PRT

; ORGANISM: Homo sapiens

; NAME/KEY: misc_feature

; LOCATION: (61)..(61)

; FEATURE:

; OTHER INFORMATION: Xaa can be any naturally occurring amino acid

; NAME/KEY: misc_feature

; LOCATION: (72)..(72)

; FEATURE:

; OTHER INFORMATION: Xaa can be any naturally occurring amino acid

US-09-724-676-52629

Query Match 40.7%; Score 37; DB 5; Length 986;

Best Local Similarity 35.0%; Pred. No. 93;

Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXXEEIIIXQETY 20
:||||: : ||||
Db 513 RNTLEKFTASIQRLIEQEEY 532

RESULT 9

US-09-724-676A-52629

; Sequence 52629, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 52629

; LENGTH: 986

; TYPE: PRT

; ORGANISM: Homo sapiens

; NAME/KEY: misc_feature

; LOCATION: (61)..(61)

; OTHER INFORMATION: Xaa can be any naturally occurring amino acid

; NAME/KEY: misc_feature

; LOCATION: (72)..(72)

; OTHER INFORMATION: Xaa can be any naturally occurring amino acid

US-09-724-676A-52629

Query Match 40.7%; Score 37; DB 5; Length 986;

Best Local Similarity 35.0%; Pred. No. 93;

Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXXEEIIIXQETY 20
:||||: : ||||
Db 513 RNTLEKFTASIQRLIEQEEY 532

RESULT 10

US-09-724-676-52622

; Sequence 52622, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 52622

; LENGTH: 1006

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676-52622

Query Match 40.7%; Score 37; DB 5; Length 1006;

Best Local Similarity 35.0%; Pred. No. 95;

Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXXEEIIIXQETY 20
:||||: : ||||
Db 533 RNTLEKFTASIQRLIEQEEY 552

RESULT 11

US-09-724-676-52635

; Sequence 52635, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 52635

; LENGTH: 1006

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676-52635

Query Match

Best Local Similarity 40.7%; Score 37; DB 5; Length 1006;

Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXXEESIIXQETY 20
: : : : :
Db 533 RNTLEKFTASIQRLIEQEEY 552

RESULT 12
US-09-724-676-52639
; Sequence 52639, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52639
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-52639

Query Match 40.7%; Score 37; DB 5; Length 1006;
Best Local Similarity 35.0%; Pred. No. 95;
Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXXEESIIXQETY 20
: : : : :
Db 533 RNTLEKFTASIQRLIEQEEY 552

RESULT 13
US-09-724-676A-52622
; Sequence 52622, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52622
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-52622

Query Match 40.7%; Score 37; DB 5; Length 1006;
Best Local Similarity 35.0%; Pred. No. 95;
Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXXEESIIXQETY 20
: : : : :
Db 533 RNTLEKFTASIQRLIEQEEY 552

RESULT 14
US-09-724-676A-52635
; Sequence 52635, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52635
; LENGTH: 1006
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-724-676A-52635

Query Match 40.7%; Score 37; DB 5; Length 1006;
Best Local Similarity 35.0%; Pred. No. 95;
Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXXEESIIXQETY 20
: : : : :
Db 533 RNTLEKFTASIQRLIEQEEY 552

RESULT 15
US-09-724-676A-52639
; Sequence 52639, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 52639
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-52639

Query Match 40.7%; Score 37; DB 5; Length 1006;
Best Local Similarity 35.0%; Pred. No. 95;
Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXXEESIIXQETY 20
: : : : :
Db 533 RNTLEKFTASIQRLIEQEEY 552

Search completed: February 11, 2003, 18:35:21
Job time : 26.0309 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:14:11 ; Search time 18.4021 Seconds
(without alignments)
109.706 Million cell updates/sec

Title: US-09-380-738A-4
Perfect score: 91
Sequence: 1 KNMEHVXXXEEIIIXQETVK 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.73: *
1: Pirl.*
2: Pirl.*
3: Pirl.*
4: Pirl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	87	95.6	222	1 KABOS2	alpha-s2-casein pr
2	59.5	65.4	235	2 A48383	alpha s2-casein
3	54.5	59.9	223	1 KASHS2	alpha-s2-casein pr
4	53.5	58.8	223	2 JN0547	alpha-s2-casein pr
5	53.5	58.8	223	2 I46995	alpha s2-casein C
6	42	46.2	566	1 HVBSU	bacillolysin (EC 3
7	40	44.0	300	2 F83804	cation efflux syst
8	39.5	43.4	333	2 E86621	aspartate dehydrog
9	39.5	43.4	333	2 G72002	aspartate dehydrog
10	39	42.9	1458	2 T51995	hypothetical prote
11	39	42.9	1458	2 T39266	probable cell cycl
12	39	42.9	2206	2 G71611	hypothetical prote
13	38	41.8	176	2 I41076	site-specific DNA-
14	38	41.8	320	2 T23904	hypothetical prote
15	38	41.8	871	2 D86355	protein T16E15.12
16	38	41.8	1025	2 E86355	hypothetical prote
17	37.5	41.2	537	2 T25536	hypothetical prote
18	37	40.7	121	2 A45278	cell division prot
19	37	40.7	121	2 G90639	cell division prot
20	37	40.7	121	2 AF0517	cell division prot
21	37	40.7	121	2 G83490	cell division prot
22	37	40.7	161	2 F86191	hypothetical prote
23	37	40.7	353	1 W2WL33	E2 protein - human
24	37	40.7	405	2 T21188	hypothetical prote
25	37	40.7	477	2 S22027	paranyosin, mini -
26	37	40.7	514	2 T48318	hypothetical prote
27	37	40.7	879	2 S22028	paranyosin, standa
28	37	40.7	888	2 H88085	protein T1P1.8 [1
29	37	40.7	1288	2 T46486	chromosomal protei

ALIGNMENTS

RESULT 1

KABOS2

alpha-s2-casein precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 01-Sep-1981 #sequence_revision 03-Feb-1994 #text_change 22-Jun-1999
C:Accession: JQ2008; A29087; A91438; S66626; A03107
R:Groenen, M.A.M.; Dijkhof, R.J.M.; Verstege, A.J.M.; van der Poel, J.J.
Gene 123, 187-193, 1993

A:Title: The complete sequence of the gene encoding bovine alphas2-casein.
A:Reference number: JQ2008; MUID:93154583; PMID:8428658

A:Accession: JQ2008

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-222 <GRO>

A:Cross-references: GB:M94327

R:Stewart, A.F.; Bonsing, J.; Beattie, C.W.; Shah, F.; Willis, I.M.; Mackinlay, A.

Mol. Biol. Evol. 4, 231-241, 1987

A:Title: Complete nucleotide sequences of bovine alpha-s2- and beta-casein cDNAs:

A:Reference number: A93062; MUID:88188989; PMID:2833669

A:Accession: A29087

A>Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-222 <STE>

A:Cross-references: GB:M16644; NID:gl62928; PIDN:AAA30479.1; PID:gl62929

R:Brignon, G.; Ribadeau Dumas, B.; Mercier, J.C.; Pelissier, J.P.; Das, B.C.

FEBS Lett. 76, 274-279, 1977

A:Title: Complete amino acid sequence of bovine alpha-s2- and beta-casein.

A:Reference number: A91438; MUID:77185633; PMID:862906

A:Contents: A allele

A:Accession: A91438

A:Molecule type: protein

A:Residues: 16-101, 'EP', 104-222 <BRI>

A>Note: four fractions, previously designated s2, s3, s4, and s6, appear to have t

these

R:Grosclaude, F.; Joudrier, P.; Mahe, M.F.

J. Dairy Res. 46, 211-213, 1979

A:Title: A genetic and biochemical analysis of a polymorphism of bovine alpha-s2-c

A:Reference number: A92771; MUID:79239837; PMID:469044

A:Contents: annotation; D allele

A>Note: the sequence of the D allele has a deletion of nine residues, which may be

R:Zucht, H.D.; Raada, M.; Adermann, K.; Maegert, H.J.; Forssmann, W.G.

FEBS Lett. 372, 185-188, 1995

A:Title: Casocidin-I: a casein-alpha(s2) derived peptide exhibits antibacterial ac

A:Reference number: S66626; MUID:9600204; PMID:7556666

A:Accession: S66626

A:Molecule type: protein

A:Residues: 163-203 <ZUC>

C:Comment: The sequence of the A allele is shown.

C:Genetics:

A:Gene: alphas2ca

A:Map position: 6

A:Introns: 17/3; 26/3; 33/3; 47/3; 56/3; 65/3; 74/3; 82/3; 97/3; 138/3; 147/3; 156

30	36	39.6	91	2	T00167
31	36	39.6	286	2	A70168
32	36	39.6	305	2	H82684
33	36	39.6	310	2	AB0312
34	36	39.6	447	2	T43750
35	36	39.6	456	2	S23104
36	36	39.6	458	2	D82974
37	36	39.6	577	2	T05136
38	36	39.6	578	2	D82902
39	36	39.6	622	2	S45129
40	36	39.6	676	2	D95249
41	36	39.6	676	2	A98114
42	36	39.6	787	2	A70132
43	36	39.6	814	2	A95206
44	36	39.6	892	2	S63199
45	36	39.6	2210	1	RRXPLC

hypothetical prote
hypothetical prote
acetyltransferase
LYSR-family transc
hypothetical prote
choline kinase - h
L-serine dehydrata
hypothetical prote
hypothetical prote
VPS27 protein - ye
transcription regu
conserved hypoteth
cell division prot
glycosyl transfera
BNi4 protein - yea
genome polypeptid

C:Superfamily: alpha-s2-casein
C:Keywords: mammary gland; milk; phosphoprotein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-223/Product: alpha-s2-casein #status experimental <MAT>
F:23,24,25,31,71,72,73,76,144,146,158/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 95.6%; Score 87; DB 1; Length 222;
Best Local Similarity 81.0%; Pred. No. 5.5e-09;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXXEESIIXQETK 21
|: ||||| ||||| |||||
Db 16 KNTMEHVSSSESIIXQETK 36
|: ||||| ||||| |||||

RESULT 2
A48383
alpha s2-casein - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 03-May-1996
C:Accession: A48383
R:Alexander, L.J.; Das Gupta, N.A.; Beattie, C.W.
Anim. Genet. 23, 365-367, 1992
A:Title: The sequence of porcine alpha s2-casein cDNA.
A:Reference number: A48383; MUID:92367960; PMID:1503276
A:Accession: A48383
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-235 <ALE>
A:Experimental source: mammary gland
A:Note: sequence inconsistent with the nucleotide translation
C:Superfamily: alpha-s2-casein

Query Match 65.4%; Score 59.5; DB 2; Length 235;
Best Local Similarity 63.6%; Pred. No. 0.0011;
Matches 14; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

Qy 1 KNTMEHVXXXEESIIXQETK 21
|: ||||| ||||| |||||
Db 16 KNTMEHVSSSESIIXQETK 37
|: ||||| ||||| |||||

RESULT 3
KASHS2
alpha-s2-casein precursor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 22-Jun-1999
C:Accession: A25070; S17856
R:Boisnard, M.; Petriassant, G.
Biochimie 67, 1043-1051, 1985
A:Title: Complete sequence of ovine alpha-s2-casein messenger RNA.
A:Reference number: A25070; MUID:86104467; PMID:3002499
A:Accession: A25070
A:Molecule type: mRNA
A:Residues: 1-223 <BOI>
A:Cross-references: GB:X03238; NID:g1238; PIDN:CAA26983.1; PID:g732894
A:Note: 64-Asn was also found
R:Boisnard, M.; Hue, D.; Bouniol, C.; Mercier, J.C.; Gaye, P.
Eur. J. Biochem. 201, 633-641, 1991
A:Title: Multiple mRNA species code for two non-allelic forms of ovine alphas2-casein.
A:Reference number: S17856; MUID:92037619; PMID:1935959
A:Accession: S17856
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-93, T, 95-223 <BO2>
C:Superfamily: alpha-s2-casein
C:Keywords: mammary gland; milk; phosphoprotein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-223/Product: alpha-s2-casein #status predicted <KA2>
F:23,24,25,32,55,72,73,74,77,145,147,159/Binding site: phosphate (Ser) (covalent) #status predicted
F:59,88,146,154,170,198/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 59.9%; Score 54.5; DB 1; Length 223;
Best Local Similarity 59.1%; Pred. No. 0.009; 7; Indels 1; Gaps 1;
Matches 13; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

Qy 1 KNTMEHVXXXEESIIXQETK 21
|: ||||| ||||| |||||
Db 16 KHKMEHVSSSEPINISQEIYK 37
|: ||||| ||||| |||||

RESULT 4
JN0547
alpha-s2-casein precursor - goat
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 31-Dec-1993 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: S33881; S33880; JN0547; S20620
R:Bouniol, C.; Brignon, G.; Mahe, M.F.; Printz, C.
Protein Seq. Data Anal. 5, 213-218, 1993
A:Title: Characterization of goat allelic alphas2-caseins A and B: further evidence
A:Reference number: S33880
A:Accession: S33881
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-223 <BOU>
A:Note: allele B
A:Accession: S33880
A:Molecule type: protein
A:Residues: 16-78, 'E', 80-223 <BO2>
A:Note: sequence deduced from compositional analysis of peptides
R:Bouniol, C.
Gene 125, 235-236, 1993
A:Title: Sequence of the goat alpha-s2-casein-encoding cDNA.
A:Reference number: JN0547; MUID:93216130; PMID:8462880
A:Accession: JN0547
A:Molecule type: mRNA
A:Residues: 1-78, 'E', 80-223 <BO3>
A:Cross-references: EMBL:X65160; NID:g955; PIDN:CAA46278.1; PID:g956
A:Note: allele A
C:Superfamily: alpha-s2-casein
C:Keywords: mammary gland; milk; phosphoprotein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-223/Product: alpha-s2-casein #status experimental <MAT>
F:23,24,25,55,72,73,74,77,145,147,159/Binding site: phosphate (Ser) (covalent) #status predicted
F:53,146/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 58.8%; Score 53.5; DB 2; Length 223;
Best Local Similarity 59.1%; Pred. No. 0.014; 7; Indels 1; Gaps 1;
Matches 13; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

Qy 1 KNTMEHVXXXEESIIXQETK 21
|: ||||| ||||| |||||
Db 16 KHKMEHVSSSEPINISQEIYK 37
|: ||||| ||||| |||||

RESULT 5
I46995
alpha s2-casein C - goat
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 13-Aug-1999
C:Accession: I46995
R:Bouniol, C.; Brignon, G.; Mahe, M.F.; Printz, C.
Anim. Genet. 25, 173-177, 1994
A:Title: Biochemical and genetic analysis of variant C of caprine alpha s2-casein
A:Reference number: I46995; MUID:95030556; PMID:7943951
A:Accession: I46995
A:Status: preliminary; translated from GB/EMBL/DDDBJ
A:Molecule type: DNA
A:Residues: 1-223 <BOU>
A:Cross-references: GB:S74171; NID:g707033; PIDN:AAB32166.1; PID:g707034
C:Superfamily: alpha-s2-casein

Query Match 58.8%; Score 53.5; DB 2; Length 223;
Best Local Similarity 59.1%; Pred. No. 0.014; 7; Indels 1; Gaps 1;
Matches 13; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

QY 1 KNTMEHVXXXEESI-IXQETKY 21
 I: ||||| :||| :||| :|||
 Db 16 KHKMEHVSSSEEPINIFQEIYK 37

RESULT 6
 HYBSU
 Bacillolysin (EC 3.4.24.28) precursor [validated] - Bacillus cereus
 C:Alternate names: Bacillus metalloendopeptidase; microbial metalloproteinase; neutral B
 C:Species: Bacillus cereus
 C:Date: 30-Jun-1988 #sequence_revision 12-Apr-1996 #text_change 15-Sep-2000
 C:Accession: S22690; A24306; I39914
 R:Wetmore, D.R.; Wong, S.L.; Roche, R.S.
 Mol. Microbiol. 6, 1593-1604, 1992
 A:Title: The role of the pro-sequence in the processing and secretion of the thermolysin
 A:Reference number: S22690; MUID:92356823; PMID:1495388
 A:Accession: S22690
 A:Molecule type: DNA
 A:Residues: 1-566 <WET>
 A:Cross-references: EMBL:M83910; NID:gl43242; PIDN:AAA22620.1; PID:gl43243
 R:Sidler, W.; Niederer, E.; Suter, F.; Zuber, H.
 Biol. Chem. Hoppe-Seyler 367, 643-657, 1986
 A:Title: The primary structure of Bacillus cereus neutral proteinase and comparison with
 A:Reference number: A24306; MUID:87000170; PMID:3092843
 A:Accession: A24306
 A:Molecule type: protein
 A:Residues: 250-566 <SID>
 A:Experimental source: strain DSM3101
 R:Stark, W.; Pauplit, R.A.; Janssonius, J.N.
 submitted to the Brookhaven Protein Data Bank, January 1992
 A:Contents: annotation: X-ray crystallography, 2.0 angstroms, residues 250-566
 A:Reference number: A31318; PDB:1NCP
 J. Mol. Biol. 199, 525-537, 1988
 A:Title: Crystal structure of neutral protease from Bacillus cereus refined at 3.0 A res
 A:Reference number: A38850; MUID:88172498; PMID:3127592
 A:Contents: annotation: X-ray crystallography, 3.0 angstroms
 C:Genetics:
 A:Gene: cnp; nprC
 C:Function:
 A:Description: hydrolyzes peptide bonds on the amino side of hydrophobic residues
 C:Superfamily: thermolysin
 C:Keywords: calcium; extracellular protein; hydrolase; metalloproteinase; zinc
 F:1-18/domain: signal sequence #status predicted <SIG>
 F:19-249/domain: activation peptide #status predicted <ACP>
 F:250-566/product: bacillolysin #status experimental <MAT>
 F:307,309,311/Binding site: calcium (Asp, Asp, Val) #status experimental
 F:388,427,435,437,440/Binding site: calcium (Asp, Glu, Glu, Glu)
 F:392,396,416/Binding site: zinc (His, His, Glu) #status experimental
 F:393,481/Active site: Glu, His #status predicted
 F:427,433,435,440/Binding site: calcium (Glu, Asn, Asp, Glu) #status experimental
 F:443,444,447,450/Binding site: calcium (Tyr, Thr, Lys, Asp) #status experimental

Query Match 46.2%; Score 42; DB 1; Length 566;
 Best Local Similarity 40.0%; Pred. No. 6.3;
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 QY 2 NTMEHVXXXEESIIXQETKY 21
 I: ||||| :||| :||| :|||
 Db 221 NTIDHVTNDKSPVKQAPK 240

RESULT 7
 FB3804
 cation efflux system (zinc/cadmium) BHL238. [imported] - Bacillus halodurans (strain C-12
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: F83804
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: F83804
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-300 <SPO>
 A:Cross-references: GB:AP001511; GB:BA000004; NID:gl0173727; PIDN:BA04957.1; GSPDB:GSPDB000004
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BHL238
 C:Superfamily: conserved hypothetical protein MJ0449

Query Match 44.0%; Score 40; DB 2; Length 300;
 Best Local Similarity 35.0%; Pred. No. 7.5;
 Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 2 NTMEHVXXXEESIIXQETKY 21
 I: ||||| :||| :||| :|||
 Db 212 NTLDHVLHDETVEMREAAK 231

RESULT 8
 E86621
 aspartate dehydrogenase [imported] - Chlamydomophila pneumoniae (strain J138)
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: E86621
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba,
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20330349; PMID:10871362
 A:Accession: E86621
 A:Molecule type: DNA
 A:Status: preliminary
 A:Residues: 1-333 <STO>
 A:Cross-references: GB:BA000008; NID:g8979421; PIDN:BAA99255.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: asd
 C:Superfamily: yeast aspartate-semialdehyde dehydrogenase

Query Match 43.4%; Score 39.5; DB 2; Length 333;
 Best Local Similarity 45.0%; Pred. No. 10;
 Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 2 NTMEHVXXXEESIIXQETKY 21
 I: ||||| :||| :||| :|||
 Db 185 NTVPHVIGEEKIL-RETVK 203

RESULT 9
 G72002
 aspartate dehydrogenase - Chlamydomophila pneumoniae (strain CWL029)
 C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
 C:Accession: G72002
 R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood
 Nature Genet. 21, 385-389, 1999
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A:Reference number: A72000; MUID:99206606; PMID:10192388
 A:Accession: G72002
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-333 <ARN>
 A:Cross-references: GB:AE001695; GB:AE001363; NID:g4377378; PIDN:AAD19185.1; PID:GSPDB:GSPDB000004
 A:Experimental source: strain CWL029
 C:Genetics:
 A:Gene: asd
 C:Superfamily: yeast aspartate-semialdehyde dehydrogenase

Query Match 43.4%; Score 39.5; DB 2; Length 333;
 Best Local Similarity 45.0%; Pred. No. 10;
 Matches 9; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 2 NTMEHVXXXEESIIXQETKY 21
 I: ||||| :||| :||| :|||

Db 185 NTVPHVGEEEKIL-RETVK 203
||:|: ||:|: ||:|

RESULT 10

T51995
hypothetical protein cut4+ - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: T51995
R:Yamashita, Y.; Nakaseko, Y.; Samejima, I.; Kumada, K.; Yamada, H.; Yanagida, M.
Nature 384, 276-279, 1996
A:Title: 20S cytosome complex formation and proteolytic activity inhibited by the cAMP/
A:Reference number: Z25896
A:Accession: T51995
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1458 <YAM>
A:Cross-references: EMBL:D85196; PIDN:BAA22618.1
C:Genetics:
A:Gene: cut4+

Query Match 42.9%; Score 39; DB 2; Length 1458;
Best Local Similarity 42.1%; Pred. No. 68;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 3 TMEHVXXXEESIIXQETK 21
||:|: ||:|: ||:|

Db 939 TMEELASNESELKNEGK 957

RESULT 11

T39266
probable cell cycle control protein cut4 [similarity] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Jun-2000
C:Accession: T39266
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Davis, P.; Churcher, C.M.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21840
A:Accession: T39266
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1458 <LYN>
A:Cross-references: EMBL:AL110295; PIDN:CAB53725.1; GSPDB:GN00067; SPDB:SPBC106.09
A:Experimental source: strain 972h-; cosmid c106
C:Genetics:
A:Gene: SPDB:SPBC106.09
A:Map position: 2

Query Match 42.9%; Score 39; DB 2; Length 1458;
Best Local Similarity 42.1%; Pred. No. 68;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 3 TMEHVXXXEESIIXQETK 21
||:|: ||:|: ||:|

Db 939 TMEELASNESELKNEGK 957

RESULT 12

G71611
hypothetical protein PFB0560w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: G71611
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.; Perle, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O. Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: G71611
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-2206 <GAR>

A:Cross-references: GB:AE001403; GB:AE001362; NID:g3845216; PIDN:AAC71901.1; PID:g3845216

A:Experimental source: clone 3D7

C:Genetics:

A:Gene: PFB0560w

Query Match 42.9%; Score 39; DB 2; Length 2206;
Best Local Similarity 36.8%; Pred. No. 11e+02;
Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 NTMEHVXXXEESIIXQETK 20
||:|: ||:|: ||:|

Db 726 NDYHHDNDVINGKLY 744

RESULT 13

141076
site-specific DNA-methyltransferase (cytosine-specific) (EC 2.1.1.73) EcoHK31I beta
C:Species: Escherichia coli
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 04-Mar-2000
C:Accession: S53984; I41076
R:Lee, K.F.; Kam, K.M.; Shaw, P.C.
Nucleic Acids Res. 23, 103-108, 1995
A:Title: A bacterial methyltransferase M.EcoHK31I requires two proteins for in vitro
A:Reference number: I41075; MUID:95175351; PMID:7870574
A:Accession: S53984
A:Molecule type: DNA
A:Residues: 1-176 <LEE>

A:Cross-references: EMBL:X82231; NID:g639971; PID:g639973

A:Experimental source: strain HK31

A:Note: the authors translated the codon TTG for residue 97 as Met

A:Note: part of this sequence, including the amino end was confirmed by protein seq

C:Complex: heterodimer; beta chain is encoded by an alternative reading frame with

C:Superfamily: Escherichia coli site-specific DNA-methyltransferase (cytosine-speci

C:Keywords: heterodimer; methyltransferase; restriction modification system; S-aden

Query Match 41.8%; Score 38; DB 2; Length 176;
Best Local Similarity 41.2%; Pred. No. 9.9;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXEESIIXQ 17
||:|: ||:|: ||:|

Db 51 QNTLEHPSVEEECTLSQ 67

RESULT 14

T23904
hypothetical protein R04D3.12 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T23904
R:Swinnburne, J.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z19815
A:Accession: T23904
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-320 <WTL>
A:Cross-references: EMBL:Z70212; PIDN:CAB54284.1; GSPDB:GN00028; CESP:R04D3.12
A:Experimental source: clone R04D3
C:Genetics:
A:Gene: CESP:R04D3.12
A:Map position: X
A:Introns: 100/3; 146/3; 215/2; 241/3

Query Match 41.8%; Score 38; DB 2; Length 320;
Best Local Similarity 39.1%; Pred. No. 19;
Matches 9; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

QY 1 KNTMEHV-XXXEESIIXQETK 21
||:|: ||:|: ||:|

Db 215 KNTLKHKNQAGNMSIACQEQYR 237

RESULT 15

D86355
protein Tl6E15.12 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D86355
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86355
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-871 <STO>
A:Cross-references: GB:AE005172; NID:99392688; PIDN:AAF87265.1; GSPDB:GN00141
C:Genetics:
A:Gene: Tl6E15.12
A:Map position: 1

Query Match 41.8%; Score 38; DB 2; Length 871;
Best Local Similarity 50.0%; Pred. No. 60;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXESI 14
| :|||||
Db 70 KKSMEHVYALEKL 83

Search completed: February 11, 2003, 18:21:17
Job time : 19.4021 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:07:35 ; Search time 9.30928 Seconds
(without alignments)
93.563 Million cell updates/sec

Title: US-09-380-738A-4
Perfect score: 91
Sequence: 1 KNTMEHVXXXESIIIXQETK 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	95.6	222	1	P02663 bos taurus
2	59.5	65.4	235	1	P39036 sus scrofa
3	54.5	59.9	223	1	P04654 ovis aries
4	53.5	58.8	223	1	P33049 capra hircu
5	42	46.2	566	1	P05806 bacillus ce
6	39	42.9	1458	1	Q9urv2 schizosacch
7	37	40.7	121	1	P22187 escherichia
8	37	40.7	353	1	P06423 human papil
9	37	40.7	477	1	P35416 drosophila
10	37	40.7	875	1	O96651 drosophila
11	37	40.7	879	1	P35790 homo sapien
12	36	39.6	456	1	P40343 saccharomyc
13	36	39.6	622	1	O34098 spiroplasma
14	36	39.6	749	1	O52177 myxococcus
15	36	39.6	757	1	P53858 saccharomyc
16	36	39.6	892	1	P14240 lymphocytic
17	36	39.6	2210	1	Q9q926 shope fibro
18	35.5	39.0	590	1	Q58690 methanococc
19	35	38.5	285	1	O00472 homo sapien
20	35	38.5	640	1	P44289 haemophilus
21	35	38.5	718	1	P36168 saccharomyc
22	35	38.5	1195	1	P04655 cavia porce
23	34.5	37.9	223	1	Q948q1 myxoma viru
24	34.5	37.9	590	1	P15026 pseudomonas
25	34	37.4	196	1	P71788 mycobacteri
26	34	37.4	265	1	O02145 lactococcus
27	34	37.4	293	1	P09575 pichia angu
28	34	37.4	367	1	Q02145 lactococcus
29	34	37.4	416	1	P36883 azotobacter
30	34	37.4	440	1	P00937 saccharomyc
31	34	37.4	484	1	Q9xec3 arabidopsis
32	34	37.4	528	1	WR42_ARATH
33	34	37.4	569	1	Y39733 MYCPN

34	34	37.4	587	1	NOT4_YEAST
35	34	37.4	612	1	YMY8_YEAST
36	34	37.4	686	1	REGG_STAAU
37	34	37.4	742	1	SYG_CABEL
38	34	37.4	1292	1	FKG2_MYCGE
39	33.5	36.8	173	1	FKG2_HUMAN
40	33.5	36.8	193	1	CAS2_CAMDR
41	33.5	36.8	267	1	XPA_CHICK
42	33.5	36.8	534	1	SR54_ASPNG
43	33	36.3	82	1	RPOH_PIRAB
44	33	36.3	82	1	RPOH_PIRHO
45	33	36.3	142	1	HBA_TARGR

ALIGNMENTS

RESULT 1	CAS2_BOVIN	STANDARD;	PRT;	222 AA.
AC	P02663; Q9TR51;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-JUL-1989 (Rel. 11, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Alpha-S2 casein precursor [Contains: Casocidin-I].			
GN	CSN1S2.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88188989; PubMed=2833669;			
RA	Stewart A.F., Bonsing J., Beattie C.W., Shah F., Willis I.M.,			
RA	Mackinlay A.G.;			
RT	"Complete nucleotide sequences of bovine alpha S2- and beta-casein			
RT	cDNAs: comparisons with related sequences in other species.";			
RL	Mol. Biol. Evol. 4:231-241(1987).			
RN	[2]			
RP	SEQUENCE OF 16-222 (A ALLELE).			
RC	TISSUE=Milk;			
RX	MEDLINE=77185633; PubMed=862906;			
RA	Brignon G., Ribadeau-Dumas B., Mercier J.-C., Pelissier J.-P.,			
RA	Das B.C.;			
RT	"Complete amino acid sequence of bovine alphas2-casein.";			
RL	FEBS Lett. 76:274-279(1977).			
RN	[3]			
RP	PARTIAL SEQUENCE (D ALLELE).			
RC	TISSUE=Milk;			
RX	MEDLINE=79239837; PubMed=469044;			
RA	Grosclaude F., Joudrier P., Mahe M.-F.;			
RT	"A genetic and biochemical analysis of a polymorphism of bovine alpha			
RT	S2-casein.";			
RL	J. Dairy Res. 46:211-213(1979).			
RN	[4]			
RP	SEQUENCE OF 165-203, AND CHARACTERIZATION OF CASOCIDIN.			
RC	TISSUE=Milk;			
RX	MEDLINE=96000204; PubMed=7556666;			
RA	Zucht H.-D., Raida M., Adermann K., Meagert H.-J., Forssmann W.-G.;			
RT	"Casocidin-1: a casein-alpha S2 derived peptide exhibits antibacterial			
RT	activity.";			
RL	FEBS Lett. 372:185-188(1995).			
CC	-I- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT			
CC	CALCIUM PHOSPHATE.			
CC	-I- FUNCTION: CASOCIDIN-I INHIBITS THE GROWTH OF ESCHERICHIA COLI AND			
CC	STAPHYLOCOCCUS CARNOSUS.			
CC	-I- SUBCELLULAR LOCATION: MAMMARY GLAND; MILK.			
CC	-I- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.			
CC	-I- MASS SPECTROMETRY: MW=4870; METHOD=Electrospray; RANGE=165-203.			
CC	-I- POLYMORPHISM: AT LEAST TWO ALLELES EXIST. THE SEQUENCE OF THE A			
CC	ALLELE IS SHOWN HERE. THE D ALLELE SEQUENCE DIFFERS FROM THAT			
CC	SHOWN IN HAVING A DELETION OF NINE RESIDUES, WHICH MAY BE 49-58,			

```

CC 50-59, OR 51-60.
CC -!- SIMILARITY: BELONGS TO THE ALPHA-CASEIN FAMILY.
CC -!- DATABASE: NAME=Protein Spotlight;
CC NOTE=Issue 16 of November 2001;
CC WWW="http://www.expasy.org/spotlight/articles/splt016.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; M16644; AAA30479.1; -.
CC DR PIR; A03107; KABOS2.
CC DR InterPro; IPR001588; Casein.
CC DR Pfam; PF00363; caseins; 2.
CC DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
CC DR Milk; Phosphorylation; Signal; Repeat; Antibiotic.
CC KW SIGNAL 1 15
CC FT CHAIN 16 222 ALPHA-S2 CASEIN.
CC FT PEPTIDE 165 203 CASOCIDIN-I.
CC FT REPEAT 76 140
CC FT REPEAT 158 222
CC FT MOD_RES 23 23 PHOSPHORYLATION.
CC FT MOD_RES 24 24 PHOSPHORYLATION.
CC FT MOD_RES 25 25 PHOSPHORYLATION.
CC FT MOD_RES 31 31 PHOSPHORYLATION.
CC FT MOD_RES 71 71 PHOSPHORYLATION.
CC FT MOD_RES 72 72 PHOSPHORYLATION.
CC FT MOD_RES 73 73 PHOSPHORYLATION.
CC FT MOD_RES 76 76 PHOSPHORYLATION.
CC FT MOD_RES 144 144 PHOSPHORYLATION.
CC FT MOD_RES 146 146 PHOSPHORYLATION.
CC FT CONFLICT 102 102 Q -> E (IN REF. 2).
CC SQ SEQUENCE 222 AA; 26019 MW; 81E7408AF1C12F7C CRC64;

Query Match 95.6%; Score 87; DB 1; Length 222;
Best Local Similarity 81.0%; Pred. No. 4.9e-09;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXEEESIIXOETVK 21
Db 16 KNTMEHVSSEESIISQETVK 36

RESULT 2
CAS2_PIG STANDARD; PRT; 235 AA.
AC P39036;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-S2 casein precursor.
GN CSNIS2.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92367960; PubMed=1503276;
RA Alexander L.J., Das Gupta N.A., Beattie C.W.;
RT "The sequence of porcine alpha s2-casein cDNA.";
RL Anm. Genet. 23:365-367(1992).
CC -!- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT
CC CALCIUM PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
CC OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92367960; PubMed=1503276;
RA Alexander L.J., Das Gupta N.A., Beattie C.W.;
RT "The sequence of porcine alpha s2-casein cDNA.";
RL Anm. Genet. 23:365-367(1992).
CC -!- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT
CC CALCIUM PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: BELONGS TO THE ALPHA-CASEIN FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X54975; CAA38719.1; -.
CC DR InterPro; IPR001588; Casein.
CC DR Pfam; PF00363; caseins; 2.
CC DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
CC DR Milk; Phosphorylation; Signal.
CC KW SIGNAL 1 15
CC FT CHAIN 16 235 ALPHA-S2 CASEIN.
CC FT MOD_RES 23 23 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 24 24 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 25 25 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 32 32 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 70 70 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 71 71 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 72 72 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 73 73 PHOSPHORYLATION (BY SIMILARITY).
CC FT MOD_RES 74 74 PHOSPHORYLATION (BY SIMILARITY).
CC SQ SEQUENCE 235 AA; 27570 MW; C903B760D184C14C CRC64;

Query Match 65.4%; Score 59.5; DB 1; Length 235;
Best Local Similarity 63.6%; Pred. No. 0.00069;
Matches 14; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 KNTMEHVXXXEEESI-IXOETVK 21
Db 16 KHEMHVSSEESIISQETVK 37

RESULT 3
CAS2_SHEEP STANDARD; PRT; 223 AA.
AC P04654;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-S2 casein precursor.
GN CSNIS2.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86104467; PubMed=3002499;
RA Boissard M., Petrisant G.;
RT "Complete sequence of ovine alpha s2-casein messenger RNA.";
RL Biochimie 67:1043-1051(1985).
CC -!- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT
CC CALCIUM PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: BELONGS TO THE ALPHA-CASEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X03238; CAA26983.1; -.
CC DR PIR; A25070; KASHS2.
CC DR InterPro; IPR001588; Casein.

```

DR Pfam; PF00363; caseins; 2.
 DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
 KW Milk; Phosphorylation; Repeat; Signal.
 FT SIGNAL 1 15 ALPHA-S2 CASEIN.
 FT CHAIN 16 223
 FT REPEAT 77 141
 FT REPEAT 159 223
 FT MOD_RES 23 23 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 24 24 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 25 25 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 72 72 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 73 73 PHOSPHORYLATION (POTENTIAL).
 FT MOD_RES 74 74 PHOSPHORYLATION (POTENTIAL).
 FT VARIANT 64 64 D -> N.
 SQ SEQUENCE 223 AA; 67212935E27426D7 CRC64;
 Query Match 59.9%; Score 54.5; DB 1; Length 223;
 Best Local Similarity 59.1%; Pred. No. 0.0056;
 Matches 13; Conservative 1; Mismatches 7; Indels 1; Gaps 1;
 QY 1 KNTMEHVXXXEST-IXOETVK 21
 I: ||||| ||| ||| |||
 Db 16 KHKMEHVSSSEEPINISQEIYK 37
 K: ||||| ||| ||| |||
 RESULT 4
 CAS2_CAPHI STANDARD; PRT; 223 AA.
 ID CAS2_CAPHI STANDARD; PRT; 223 AA.
 AC P33049;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-S2 casein precursor (Alpha-S2-CN).
 GN CSN1S2.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93216130; PubMed=8462880;
 RA Bouniol C.;
 RT "Sequence of the goat alpha s2-casein-encoding cDNA."
 RL Gene 125:235-236(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bouniol C., Brignon G., Mahe M.-F., Printz C.;
 RT "Characterization of goat allelic alpha-s2-caseins A and B: further
 RT evidence of the phosphorylation code of caseins.";
 RL Protein Seq. Data Anal. 5:213-218(1993).
 RN [3]
 RP SEQUENCE FROM N.A. (VARIANT C).
 RX MEDLINE=95030556; PubMed=7943951;
 RA Bouniol C., Brignon G., Mahe M.F., Printz C.;
 RT "Biochemical and genetic analysis of variant C of caprine alpha s2-
 RT casein (Capra hircus)."
 RL Anim. Genet. 25:173-177(1994).
 CC -1- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT
 CC CALCIUM PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
 CC -1- POLYMORPHISM: THREE ALLELES OF ALPHA-S2 CASEIN ARE KNOWN: A, B AND
 CC C. THE FREQUENCIES OF THE ALLELES IS ESTIMATED TO BE 0.85, 0.04
 CC AND 0.11 IN THE FRENCH DAIRY BREEDS 'ALPINE' AND 'SAANEN'.
 CC -1- SIMILARITY: BELONGS TO THE ALPHA-CASEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC
 DR EMBL; X65160; CAA46278.1; -
 DR EMBL; S74171; AAB32166.1; -
 DR PIR; S20620; S20620.
 DR PIR; JN0547; JN0547.
 DR PIR; S33880; S33880.
 DR PIR; S33881; S33881.
 DR InterPro; IPR001588; Casein.
 DR Pfam; PF00363; caseins; 2.
 DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
 KW Milk; Phosphorylation; Signal; Repeat; Polymorphism.
 FT SIGNAL 1 15 BY SIMILARITY.
 FT CHAIN 16 223 ALPHA-S2 CASEIN.
 FT REPEAT 77 141
 FT REPEAT 159 223
 FT MOD_RES 23 23 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 24 24 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 25 25 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 72 72 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 73 73 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 74 74 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 77 77 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 145 145 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 147 147 PHOSPHORYLATION (BY SIMILARITY).
 FT VARIANT 79 79 E -> K (IN VARIANT B).
 FT VARIANT 182 182 K -> I (IN VARIANT C).
 SQ SEQUENCE 223 AA; 187DEF42FD688291 CRC64;
 Query Match 58.8%; Score 53.5; DB 1; Length 223;
 Best Local Similarity 59.1%; Pred. No. 0.0086;
 Matches 13; Conservative 1; Mismatches 7; Indels 1; Gaps 1;
 QY 1 KNTMEHVXXXEST-IXOETVK 21
 I: ||||| ||| ||| |||
 Db 16 KHKMEHVSSSEEPINIFQEIYK 37
 K: ||||| ||| ||| |||
 RESULT 5
 NPPE_BACCE STANDARD; PRT; 566 AA.
 ID NPPE_BACCE STANDARD; PRT; 566 AA.
 AC P05806;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bacillolysin precursor (EC 3.4.24.28) (Neutral protease).
 GN NPR OR NPRC.
 OS Bacillus cereus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1396;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92356823; PubMed=1495388;
 RA Wetmore D.R., Wong S.L., Roche R.S.;
 RT "The role of the pro-sequence in the processing and secretion of the
 RT thermolysin-like neutral protease from Bacillus cereus.";
 RL Mol. Microbiol. 6:1593-1604(1992).
 RN [2]
 RP SEQUENCE OF 250-566.
 RC STRAIN=DSM 3101;
 RX MEDLINE=87000170; PubMed=3092843;
 RA Sidler W., Niederer E., Suter F., Zuber H.;
 RT "The primary structure of Bacillus cereus neutral proteinase and
 RT comparison with thermolysin and Bacillus subtilis neutral
 RT proteinase.";
 RL Biol. Chem. Hoppe-Seyler 367:643-657(1986).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE=88172498; PubMed=3127592;
 RA Pauptit R.A., Karlsson R., Picot D., Jenkins J.A.,
 RA Niklaus-Reimer A.-S., Jansonsius J.N.;
 RT "Crystal structure of neutral protease from Bacillus cereus refined
 RT at 3.0-A resolution and comparison with the homologous but more

thermostable enzyme thermolysin."; [4]
 J. Mol. Biol. 199:525-537(1988).
 X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS). MEDLINE=92339470; PubMed=1633827;
 Stark W., Paupit R.A., Wilson K.S., Jansonius J.N.; "The structure of neutral protease from *Bacillus cereus* at 0.2-nm resolution."; Eur. J. Biochem. 207:781-791(1992). [5]
 X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF MUTANT SER-393. Lister S.A., Wetmore D.R., Roche R.S., Coddling P.W.; "E144S active-site mutant of the *Bacillus cereus* thermolysin-like neutral protease at 2.8-A resolution."; Acta Crystallogr. D 52:543-550(1996).
 -1- FUNCTION: THERMOLABILE EXTRACELLULAR ZINC METALLOPROTEASE. CC
 -1- CATALYTIC ACTIVITY: Similar, but not identical, to that of thermolysin. CC
 -1- COFACTOR: BINDS 1 ZINC ION. BINDS FOUR CALCIUM IONS (BY SIMILARITY). CC
 -1- SUBCELLULAR LOCATION: Secreted. CC
 -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4. CC

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL: M83910; AAA22620.1; --
 PIR: A24306; HYBSU.
 PDB: INPC; 31-OCT-93.
 PDB: IESP; 07-DEC-95.
 MEROPS: M04.015; --
 InterPro: IPR005075; Pep_M4_propep.
 InterPro: IPR001570; Peptidase_M4.
 InterPro: IPR000130; Zn_MTPeptidse.
 Pfam: PF01447; Peptidase_M4; 1.
 Pfam: PF02868; Peptidase_M4_C; 1.
 Pfam: PF03413; Pep_M4_propep; 1.
 PRINTS: PR00730; THERMOLYSIN.
 PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Hydrolyase; Metalloprotease; Calcium; zinc; 3D-structure; Signal.
 FT SIGNAL 1 27
 FT PROPEP 28 249
 FT CHAIN 250 566
 FT METAL 392 392
 FT ACT_SITE 393 393
 FT METAL 396 396
 FT METAL 416 416
 FT ACT_SITE 481 481
 FT STRAND 253 260
 FT TURN 262 263
 FT STRAND 265 274
 FT TURN 275 276
 FT STRAND 277 279
 FT STRAND 281 282
 FT STRAND 289 297
 FT STRAND 303 304
 FT STRAND 306 307
 FT STRAND 311 312
 FT HELIX 315 338
 FT TURN 342 345
 FT STRAND 350 356
 FT STRAND 363 365
 FT STRAND 370 372
 FT STRAND 380 380
 FT HELIX 383 385
 FT HELIX 387 400
 FT TURN 401 403
 FT HELIX 409 429

FT TURN 430 431
 FT STRAND 437 438
 FT TURN 440 442
 FT STRAND 443 443
 FT TURN 445 446
 FT STRAND 453 454
 FT HELIX 458 461
 FT TURN 462 462
 FT HELIX 467 469
 FT TURN 475 496
 FT STRAND 498 500
 FT TURN 501 502
 FT STRAND 503 505
 FT TURN 510 523
 FT TURN 524 524
 FT TURN 527 528
 FT HELIX 531 546
 FT TURN 548 549
 FT HELIX 551 562
 FT TURN 563 564
 SQ SEQUENCE 566 AA; 60919 MW; E18B4572C2C4E1D3 CRC64;
 Query Match 46.2%; Score 42; DB 1; Length 566;
 Best Local Similarity 40.0%; Pred.No. 3.2;
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 Qy 2 NTMEHVXXXEEIIIXQETK 21
 ||:||||:|:|:|
 Db 221 NTIDHVTNDKSPVKQEPK 240
 RESULT 6
 CUT4_SCHPO STANDARD; PRT: 1458 AA.
 AC Q9URV2: O13457;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cut4 protein.
 GN CUT4 OR SPBC106.09.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RX MEDLINE=97078755; PubMed=8918880;
 RA Yamashita Y.M., Nakaseko Y., Samejima I., Kumada K., Yamada H.,
 Michaelson D., Yanagida M.;
 RT "20S cyclosome complex formation and proteolytic activity inhibited by the CAMP/PKA pathway."; Nature 384:276-279(1996). [2]
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B., Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Calibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Has an essential role in the regulation of 20S cyclosome
CC complex formation. Mutations to this protein prevent the exit from
CC mitosis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D85196; BAA22618.1; -;
CC EMBL; AL110295; CAB53725.1; -;
CC Mitosis.
CC CONFLICT 995 995 A -> V (IN REF. 1).
CC SEQUENCE 1458 AA; 165409 MW; 182E65BDA3A82183 CRC64;
CC
CC Query Match 42.9%; Score 39; DB 1; Length 1458;
CC Best Local Similarity 42.1%; Pred. No. 31;
CC Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
CC
CC QY 3 TMEHVXXEESIIXQETK 21
CC ||| : || : |||
CC Db 939 TMEILASNESEKNEGK 957
CC
CC RESULT 7
CC FTSL_ECOLI STANDARD; PRT; 121 AA.
CC ID FTSL_ECOLI
CC AC P22187;
CC DT 01-AUG-1991 (Rel. 19, Created)
CC DT 01-DEC-1992 (Rel. 24, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Cell division protein ftsl.
CC GN FTSL OR MRAR OR B0083 OR Z0093 OR ECS0087.
CC OS Escherichia coli, and
CC OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC OC Escherichia.
CC OX NCBI_TaxID=562, 83334;
CC [J]
CC RN SEQUENCE FROM N.A.
CC RX MEDLINE=93077455; PubMed=1332942;
CC RA Guzman L.M., Barondess J.J., Beckwith J.;
CC "ftsl, an essential cytoplasmic membrane protein involved in cell
CC division in Escherichia coli."
CC RL J. Bacteriol. 174:7716-7728(1992).
CC [2]
CC RN SEQUENCE FROM N.A.
CC RX MEDLINE=93077472; PubMed=1447153;
CC RA Ueki M., Wachi M., Jung H.K., Ishino F., Matsuhashi M.;
CC "Escherichia coli mrar gene involved in cell growth and division."
CC RL J. Bacteriol. 174:7841-7843(1992).
CC [3]
CC RN SEQUENCE FROM N.A.
CC RX STRAIN=K12 / MG1655;
CC RX MEDLINE=97426617; PubMed=9278503;
CC RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
CC RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
CC RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
CC RA Mau B., Shao Y.;
CC "The complete genome sequence of Escherichia coli K-12.";

Science 277:1453-1474(1997).
[4]
RN PRELIMINARY SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=83296957; PubMed=6350821;
RA Nakamura M., Maruyama I.N., Soma M., Kato J., Suzuki H., Horota Y.;
RT "On the process of cellular division in Escherichia coli: nucleotide
RT sequence of the gene for penicillin-binding protein 3.";
RL Mol. Gen. Genet. 191:1-9(1983).
[5]
RN SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
[6]
RN SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 050952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001).
[7]
RN SEQUENCE OF 1-9 FROM N.A.
RC STRAIN=K12 / W3110;
RX MEDLINE=90251464; PubMed=2187182;
RA Gomez M.J., Fluoret B., van Heijenoort J., Ayala J.A.;
RT "Nucleotide sequence of the regulatory region of the gene pbpB of
RT Escherichia coli.";
RL Nucleic Acids Res. 18:2813-2813(1990).
CC -!- FUNCTION: PROTEIN INVOLVED IN CELL DIVISION AND CELL GROWTH. MAY
CC PLAY SOME ROLE IN COUPLING CELL DIVISION AND PEPTIDOGLYCAN
CC PHYSIOLOGY.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
CC -!- SIMILARITY: BELONGS TO THE FTSL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S49875; AAB24309.1; -;
CC EMBL; S49802; AAB24311.1; -;
CC EMBL; X55034; CAA38860.1; -;
CC EMBL; AE000118; AAC73194.1; -;
CC EMBL; K00137; AAA24299.1; -;
CC EMBL; AE005185; AAG54387.1; -;
CC EMBL; AP002550; BAB33510.1; -;
CC PIR; S14387; S14387.
CC PIR; A45278; A45278.
CC EcoGene; EG11086; ftsl.
CC Cell division; Transmembrane; Inner membrane; Complete proteome.
CC DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 38 57 POTENTIAL.
CC DOMAIN 58 121 PERIPLASMIC (POTENTIAL).
CC SEQUENCE 121 AA; 13627 MW; 60D5DF7B2ECB7090 CRC64;
CC
CC Query Match 40.7%; Score 37; DB 1; Length 121;
CC Best Local Similarity 38.9%; Pred. No. 5.3;

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X62591; CAA44476.1; -;
 DR EMBL; AE003554; AAF50371.1; -;
 DR PIR: S22027; S22027.
 DR FlyBase; FBgn0003149; Pm.
 DR InterPro; IPR002928; Myosin_tail.
 KW Pfam; PF01576; Myosin_tail; 1.
 KW Coiled coil; Muscle protein; Thick filament; Myosin; Phosphorylation;
 KW Alternative splicing.
 FT DOMAIN 1 107 NONHELICAL REGION (POTENTIAL).
 FT DOMAIN 108 456 COILED COIL (POTENTIAL).
 FT DOMAIN 457 477 NONHELICAL REGION (POTENTIAL).
 SQ SEQUENCE 477 AA; 54889 MW; 3FD4876F12EB0828 CRC64;

Query Match 40.7%; Score 37; DB 1; Length 477;
 Best Local Similarity 38.1%; Pred. No. 23;
 Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXEESIIXQETVK 21
 I: | | : | | : | |
 Db 291 KHVVEQVHEQERIVKLETIK 311

RESULT 10
 TP3B_DROME
 ID TP3B_DROME STANDARD; PRT; 875 AA.
 AC O96651; Q9W416;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA topoisomerase III beta (EC 5.99.1.2).
 OS TOP3-BETA OR TOP3 OR CG3458.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

[1]
 RP SEQUENCE FROM N.A., FUNCTION, AND DEVELOPMENTAL STAGE.
 RC TISSUE=Embryo;
 RX MEDLINE=20102653; PubMed=10636841;
 RA Wilson T.M., Chen A.D., Hsieh T.-S.;
 RT "Cloning and characterization of Drosophila topoisomerase IIibeta.
 RL Relaxation of hypernegatively supercoiled DNA.";
 RL J. Biol. Chem. 275:1533-1540(2000).

[2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -I- FUNCTION: WEAKLY RELAXES NEGATIVE SUPERCOILS AND DISPLAYS A
 CC DISTINCT PREFERENCE FOR BINDING SINGLE-STRANDED DNA.
 CC -I- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
 CC DNA, followed by passage and rejoining.
 CC -I- DEVELOPMENTAL STAGE: EXPRESSED DURING THE FIRST 6 HOURS OF
 CC EMBRYONIC DEVELOPMENT, LEVELS DECLINE DURING LARVAL AND PUPAL
 CC STAGES TO INCREASE AGAIN DURING ADULTHOOD.
 CC -I- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA
 CC BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN
 CC WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS
 CC AT ONE END OF THE ENZYME-SEVERED DNA STRAND (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE
 CC FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; AF099909; AAD13219.1; -;
 DR EMBL; AE003437; AAF46144.1; -;
 DR FlyBase; FBgn0026015; Top3-beta.
 DR InterPro; IPR002936; DNATprim_Toprim.
 DR InterPro; IPR003601; DNATopI_ATP_bind.
 DR InterPro; IPR003602; DNATopI_DNA_bind.
 DR InterPro; IPR000380; Prok_Tpismrase.
 DR Pfam; PF01131; Topoisom_bac; 1.
 DR Pfam; PF01751; Toprim; 1.
 DR SMART; SM00437; TOPIAC; 1.
 DR SMART; SM00436; TOPIBC; 1.
 DR SMART; SM00493; TOPRIM; 1.
 DR SMART; SM00493; TOPRIM; 1.
 DR PROSITE; PS00396; TOPOISOMERASE_I_PROK; 1.
 KW Isomerase; Topoisomerase; DNA-binding.
 FT ACT_SITE 332 332 DNA CLEAVAGE (BY SIMILARITY).
 FT DOMAIN 819 859 GLY-RICH.
 FT CONFLICT 747 747 V -> M (IN REF. 1).
 SQ SEQUENCE 875 AA; 96973 MW; 3A26520C10AB6057 CRC64;

Query Match 40.7%; Score 37; DB 1; Length 875;
 Best Local Similarity 41.2%; Pred. No. 43;
 Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 5 EHVXXEESIIXQETVK 21
 I: | | : | | : | |
 Db 275 EHKATVESVSKAYK 291

RESULT 11
 MISP_DROME
 ID MISP_DROME STANDARD; PRT; 879 AA.
 AC P35415; Q9VSP6;


```

RT paramyosin.";
RL J. Mol. Biol. 220:687-700(1991).
CC -!- FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF MANY THICK
CC FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.
CC -!- SUBUNIT: HETRODIMER OF TWO ISOFORMS (POTENTIAL).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; A LONG FORM (SHOWN
CC HERE) AND A SHORT FORM (AC P35416); ARE PRODUCED BY ALTERNATIVE
CC SPLICING AND/OR POST-TRANSLATIONAL MODIFICATIONS.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL LARVAL AND ADULT MUSCLE
CC TISSUES. EXPRESSION IS FIVE TIMES HIGHER IN TUBULAR THAN IN
CC FIBRILLAR MUSCLES.
CC -!- DEVELOPMENTAL STAGE: UNDETECTABLE DURING GASTRULATION AND EARLY
CC PHASES OF GERM BAND FORMATION. INCREASES DURING ORGANOGENESIS,
CC AROUND 10 HOURS POSTFERTILIZATION, TO THE ADULT STAGE.
CC -!- PTM: THE MORE ACIDIC AND LESS-ABUNDANT ISOFORM IS PHOSPHORYLATED
CC IN VIVO.
CC -!- SIMILARITY: HIGH, TO MYOSIN HEAVY CHAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X62590; CAA44475.1; -.
DR EMBL; X58722; CAA41557.1; -.
DR EMBL; AE003554; AAF50370.1; -.
DR PIR; S20486; S20486.
DR PIR; S22028; S22028.
DR FlyBase; FBgn0003149; Ptm.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF01576; Myosin_tail; 1.
DR Coiled coil; Muscle protein; Thick filament; Myosin; Phosphorylation;
DR KW Alternative splicing.
DR FT DOMAIN 1 31 NONHELICAL REGION (POTENTIAL).
DR FT DOMAIN 32 858 COILED COIL (POTENTIAL).
DR FT DOMAIN 859 879 NONHELICAL REGION (POTENTIAL).
DR FT DISULFID 368 368 INTERCHAIN (POTENTIAL).
DR FT DISULFID 784 784 INTERCHAIN (POTENTIAL).
DR FT CONFLICT 500 500 MISSING (IN REF. 2).
DR SQ SEQUENCE 879 AA; 102338 MW; 85BB333519815A1A CRC64;

Query Match 40.7%; Score 37; DB 1; Length 879;
Best Local Similarity 38.1%; Pred. No. 43;
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXEESIIXQETVK 21
   |: :|| :|| :|| :||
Db 693 KHVVEQVHEEGRIVKLETK 713

RESULT 12
KICH_HUMAN
ID KICH_HUMAN STANDARD; PRT; 456 AA.
AC P35790;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Choline kinase (EC 2.7.1.32) (CK) (CHETK-alpha).
GN CHK OR CKI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92316236; PubMed=1618328;
RA Hosaka K., Tanaka S., Nikawa J.-I., Yamashita S.;
RT "Cloning of a human choline kinase cDNA by complementation of the
RT yeast cki mutation.";
```


RL FEBS Lett. 304:229-232(1992).
CC -!- FUNCTION: MAY HAVE A REGULATORY ROLE IN PHOSPHATIDYLCHOLINE
CC SYNTHESIS.
CC -!- CATALYTIC ACTIVITY: ATP + choline = ADP + O-phosphocholine.
CC -!- PATHWAY: INITIAL STEP OF THE CDP-CHOLINE AND CDP-ETHANOLAMINE
CC PATHWAYS.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE CHOLINE/ETHANOLAMINE KINASES FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D10704; BAA01547.1; -
CC PIR: S23104; S23104.
CC Genew: HGNC:1937; CHK.
CC MIM: 118491; -
CC InterPro: IPR002573; Choline_kinase.
CC Pfam: PF01633; Choline_kinase; 1.
CC Transferrase; Kinase.
KW DOMAIN 50 84 PRO-RICH.
FT ACT_SITE 303 303 BY SIMILARITY.
SQ SEQUENCE 456 AA; 52065 MW; BD8D13D102178E97 CRC64;

Query Match 39.6%; Score 36; DB 1; Length 456;
Best Local Similarity 38.9%; Pred. No. 33;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 KNTMEHVXXXFEESIIXQ 18
Db 386 QNDFFNLSTEKSIREE 403

RESULT 13
VP27_YEAST
ID VP27_YEAST STANDARD; PRT; 622 AA.
AC P40343;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Vacuolar protein sorting-associated protein VPS27.
GN VPS27 OR GRD11 OR YNR006W OR N2038.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96042307; PubMed=7593183;
RA Piper R.C., Cooper A.A., Yang H., Stevens T.H.;
RT "VPS27 controls vacuolar and endocytic traffic through a prevacuolar
RT compartment in Saccharomyces cerevisiae.";
RL J. Cell Biol. 131:603-617(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96042307; PubMed=7593183;
RA Piper R.C., Cooper A.A., Yang H., Stevens T.H.;
RT "VPS27 controls vacuolar and endocytic traffic through a prevacuolar
RT compartment in Saccharomyces cerevisiae.";
RL J. Cell Biol. 131:603-617(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96042307; PubMed=7593183;
RA Verhasselt P., Aert R., Voet M., Volckaert G.;
RT "Twelve open reading frames revealed in the 23.6 kb segment flanking
RT the centromere on the Saccharomyces cerevisiae chromosome XIV right
RT arm.";
RL Yeast 10:1355-1361(1994).
CC -!- FUNCTION: REQUIRED FOR MEMBRANE TRAFFIC TO THE VACUOLE.
CC -!- SIMILARITY: CONTAINS 1 FYVE-TYPE ZINC FINGER.
CC -!- SIMILARITY: CONTAINS 1 VHS DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
EMBL: U24218; AAA96002.1; -
EMBL: X77395; CAA54574.1; -
EMBL: Z71620; CAA96282.1; -
PIR: S45129; S45129
SGD: S0005289; VPS27.
InterPro: IPR002014; HRS.
InterPro: IPR002950; Josephin.
InterPro: IPR003903; UIM.
InterPro: IPR003036; Znf_FYVE.
Pfam: PF00790; VHS; 1.
Pfam: PF01363; FYVE; 1.
Pfam: PF02809; UIM; 2.
PRODom: PD003686; HRS; 1.
SMART: SM00064; FYVE; 1.
SMART: SM00288; VHS; 1.
PROSITE: PS0178; ZF_FYVE; 1.
PROSITE: PS0179; VHS; 1.
Zinc-finger. 18 149 VHS.
FT ZN_FING 170 230 FYVE-TYPE.
FT CONFLICT 321 322 NV -> KL (IN REF. 2).
SQ SEQUENCE 622 AA; 70945 MW; 022C23CBAB2E7E1D CRC64;

Query Match 39.6%; Score 36; DB 1; Length 622;
Best Local Similarity 58.3%; Pred. No. 46;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 NTMEHVXXXFEES 13
Db 95 DTMEHVILREDS 106

RESULT 14
SPOT_SPICI
ID SPOT_SPICI STANDARD; PRT; 749 AA.
AC Q34098;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (EC 3.1.7.2)
DE ((ppGppase) (Penta-phosphate guanosine-3'-pyrophosphohydrolase).
GN SPOT.
OS Spiroplasma citri.
OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
OC Spiroplasmataceae; Spiroplasma.
OX NCBI_TaxID=2133;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GI-3;
RX MEDLINE=97386419; PubMed=9244268;
RA Jacob C., Nouzieres F., Duret S., Bove J.M., Renaudin J.;
RT "Isolation, characterization, and complementation of a motility
RT mutant of Spiroplasma citri.";
RL J. Bacteriol. 179:4802-4810(1997).
CC -!- FUNCTION: IN EUBACTERIA PPGBP (GUANOSINE 3'-DIPHOSPHATE 5'-
CC DIPHOSPHATE) IS A MEDIATOR OF THE STRINGENT RESPONSE THAT
CC COORDINATES A VARIETY OF CELLULAR ACTIVITIES IN RESPONSE TO
CC CHANGES IN NUTRITIONAL ABUNDANCE. THIS ENZYME CATALYSES THE
CC DEGRADATION OF PPGBP INTO GDP. IT MAY ALSO BE CAPABLE OF
CC CATALYZING THE SYNTHESIS OF PPGBP (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Guanosine 3',5'-bis(diphosphate) + H(2)O =
CC guanosine 5'-diphosphate + diphosphate.
CC -!- COFACTOR: MANGANESE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE RELA / SPOT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

theCtftio, tFk003007, ME_Fpase_HDC:

```

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:13:41 ; Search time 23.5979 Seconds
(without alignments)
183.363 Million cell updates/sec

Title: US-09-380-738a-4
Perfect score: 91
Sequence: 1 KNTMEHVXXXESIIIXQETK 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriaph.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	90.1	210	6	O62825 bubalus bub
2	53.5	58.8	124	6	Q9MYU7 capra hircu
3	53.5	58.8	223	6	Q9TTQ7 capra hircu
4	53.5	58.8	223	6	Q9GK07 capra hircu
5	52.5	57.7	223	6	Q9MYU6 capra hircu
6	43	47.3	953	4	Q9ULD6 homo sapien
7	42	46.2	566	2	O32309 bacillus th
8	40.5	44.5	334	16	Q9PK31 chlamydia m
9	40	44.0	300	16	Q9KDH5 bacillus ha
10	39.5	43.4	333	16	Q9JS73 chlamydia p
11	39.5	43.4	333	16	Q9Z6L1 chlamydia p
12	39	42.9	235	9	Q8SCS9 pseudomonas
13	39	42.9	2206	5	O96205 plasmodium
14	38.5	42.3	2647	5	Q9UX0 plasmodium
15	38	41.8	176	2	O52510 enterobacte
16	38	41.8	176	2	Q47257 escherichia

17	38	41.8	203	2	Q93FK1	Q93fk1 citrobacter
18	38	41.8	217	1	O97YK6	Q977K6 uncultured
19	38	41.8	242	5	Q9VYM8	Q9Vym8 drosophila
20	38	41.8	320	5	Q9U399	Q9u399 caenorhabdi
21	38	41.8	396	2	Q59719	Q59719 pseudomonas
22	38	41.8	503	11	Q9ERZ0	Q9erz0 mus musculu
23	38	41.8	871	10	Q9LME2	Q9lme2 arabidopsis
24	38	41.8	891	4	Q9Y2K3	Q9y2k3 homo sapien
25	38	41.8	1025	10	Q9LME3	Q9lme3 arabidopsis
26	37.5	41.2	537	5	P91038	P91038 caenorhabdi
27	37	40.7	92	12	Q68395	Q68395 human cytom
28	37	40.7	121	16	Q8XEN0	Q8xen0 salmoneilla
29	37	40.7	149	10	Q8RV68	Q8rv68 arabidopsis
30	37	40.7	161	10	Q9SYL4	Q9syl4 arabidopsis
31	37	40.7	312	2	Q93RA5	Q93ra5 thermus the
32	37	40.7	405	5	Q19671	Q19671 caenorhabdi
33	37	40.7	514	10	Q9LYI8	Q9lyy8 arabidopsis
34	37	40.7	566	2	Q48857	Q48857 lactobacill
35	37	40.7	566	2	Q9AQ52	Q9aq52 bacillus th
36	37	40.7	888	5	Q9U8C7	Q9u8c7 caenorhabdi
37	37	40.7	1175	10	Q9FJK1	Q9fjk1 arabidopsis
38	37	40.7	1175	10	Q9C5Y4	Q9c5y4 arabidopsis
39	37	40.7	1177	10	Q9FUY9	Q9fuy9 arabidopsis
40	37	40.7	1202	4	Q9UNT9	Q9unt9 homo sapien
41	37	40.7	1288	4	O95752	O95752 homo sapien
42	37	40.7	1288	4	O9NTJ3	O9ntj3 homo sapien
43	37	40.7	1308	5	O8T6H2	O8t6h2 dictyostelli
44	37	40.7	1896	5	Q9U8M0	Q9u8m0 periplaneta
45	36	39.6	91	9	O80072	O80072 staphylococ

ALIGNMENTS

RESULT 1

O62825 PRELIMINARY; PRT; 210 AA.
 AC O62825;
 DT 01-AUG-1998 (TRENBLrel. 07, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
 DE AS2-casein (Fragment).
 OS Bubalus bubalis (Domestic water buffalo).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bubalus.
 OX NCBI_TaxID=89462;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MAMMARY GLAND;
 RA Das P., Jain S., Garq L.C.;
 RT "Cloning and nucleotide sequence of cDNA encoding aS2-casein in B.
 RL bubalis.";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ005431; CAA06534.2; -
 DR InterPro: IPR001588; Casein.
 DR Pfam: PF00363; caseins; 2.
 FT NON_TER
 SQ SEQUENCE 210 AA; 24700 MW; 05DEF95963F1132C CRC64;

Query Match 90.1%; Score 82; DB 6; Length 210;
 Best Local Similarity 76.2%; Pred. No. 1.2e-07;
 Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KNTMEHVXXXESIIIXQETK 21
 I:||||| ||||| |||||
 Db 4 KHTMEHVSSSESIISQETK 24

RESULT 2

Q9MYU7 PRELIMINARY; PRT; 124 AA.
 ID Q9MYU7
 AC Q9MYU7;

DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE Alpha s2-casein.
 GN CSN1S2.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lagonigro R., Pilla F., Matassino D., Zullo A.;
 RT "Sequence of goat alpha s2-casein allele 0 encoding cDNA."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ289715; CAB94235.1; -
 DR EMBL; AJ289715; CAB94235.1; -
 DR InterPro; IPR001588; Casein.
 DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
 SQ SEQUENCE 124 AA; 14533 MW; C363E536CC17B5F2 CRC64;

 Query Match 58.8%; Score 53.5; DB 6; Length 124;
 Best Local Similarity 59.1%; Pred. No. 0.018;
 Matches 13; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

 QY 1 KNTMEHVXXXEESI-IXQETK 21
 Db 16 KHKMEHVSSSEEPINIFQEIYK 37

 RESULT 3
 Q9TTQ7 PRELIMINARY; PRT; 223 AA.
 AC Q9TTQ7;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE Alpha s2-casein.
 GN CSN1S2.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Veltri C.C., Pilla F.F., Lagonigro R.R.;
 RT "A new allele of goat alpha s2-casein."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ249995; CAB59920.1; -
 DR InterPro; IPR001588; Casein.
 DR Pfam; PF00363; caseins; 2.
 DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
 SQ SEQUENCE 223 AA; 26433 MW; CE9F4DC8D768B293 CRC64;

 Query Match 58.8%; Score 53.5; DB 6; Length 223;
 Best Local Similarity 59.1%; Pred. No. 0.034;
 Matches 13; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

 QY 1 KNTMEHVXXXEESI-IXQETK 21
 Db 16 KHKMEHVSSSEEPINIFQEIYK 37

 RESULT 4
 Q9GK07 PRELIMINARY; PRT; 223 AA.
 AC Q9GK07;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Alpha s2-casein.
 GN CSN1S2.
 OS Capra hircus (Goat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Veltri C., Pilla F., Lagonigro R.;
 RT "A new allele of alpha s2-casein."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ297310; CAC21704.2; -
 DR EMBL; AJ297311; CAC21704.2; JOINED.
 DR EMBL; AJ242728; CAC21704.2; JOINED.
 DR EMBL; AJ297312; CAC21704.2; JOINED.
 DR EMBL; AJ297313; CAC21704.2; JOINED.
 DR EMBL; AJ297314; CAC21704.2; JOINED.
 DR EMBL; AJ242527; CAC21704.2; JOINED.
 DR EMBL; AJ297315; CAC21704.2; JOINED.
 DR EMBL; AJ297316; CAC21704.2; JOINED.
 DR EMBL; AJ242528; CAC21704.2; JOINED.
 DR EMBL; AJ242528; CAC21704.2; JOINED.
 DR EMBL; AJ242533; CAC21704.2; JOINED.
 DR InterPro; IPR001588; Casein.
 DR Pfam; PF00363; caseins; 2.
 DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 2.
 SQ SEQUENCE 223 AA; 26432 MW; CE9765BBD7688C9D CRC64;

 Query Match 58.8%; Score 53.5; DB 6; Length 223;
 Best Local Similarity 59.1%; Pred. No. 0.034;
 Matches 13; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

 QY 1 KNTMEHVXXXEESI-IXQETK 21
 Db 16 KHKMEHVSSSEEPINIFQEIYK 37

 RESULT 5
 Q9MYU6 PRELIMINARY; PRT; 223 AA.
 AC Q9MYU6;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE Alpha s2-casein.
 GN CSN1S2.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxID=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lagonigro R., Pilla F., Matassino D., Zullo A.;
 RT "A new allele of goat alpha s2-casein gene."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ289716; CAB94236.1; -
 DR InterPro; IPR001588; Casein.
 DR Pfam; PF00363; caseins; 2.
 DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
 SQ SEQUENCE 223 AA; 26403 MW; OE1FE83F24DA85E2 CRC64;

 Query Match 57.7%; Score 52.5; DB 6; Length 223;
 Best Local Similarity 54.5%; Pred. No. 0.053;
 Matches 12; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

 QY 1 KNTMEHVXXXEESI-IXQETK 21
 Db 16 KHKMEHVSSSEEPINIFQEIYK 37

 RESULT 6
 Q9ULD6 PRELIMINARY; PRT; 953 AA.
 ID Q9ULD6
 AC Q9ULD6;

Qy	2	NTMEHVXXXESIIIXQETKYK	21
		: :	
Dd	221	NTIDHVTNDKSPVKQEAPK	240
RESULT 8			
Q9PK31			
ID	Q9PK31	PRELIMINARY;	PRT; 334 AA.
AC	Q9PK31;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Aspartate semialdehyde dehydrogenase.		
GN	TC0642.		
OS	Chlamydia muridarum.		
OC	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.		
OX	NCBI_TaxID=83560;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MOPN / NIGG;		
RX	MEDLINE=20150255; PubMed=10684935;		
RA	Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,		
RA	White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,		
RA	Linhner K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,		
RA	Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.		
EI	Eisen J., Fraser C.M.;		
RT	"Genome sequences of Chlamydia trachomatis MoPn and Chlamydia		
RL	pneumoniae AR39.";		
DR	Nucleic Acids Res. 28:1397-1406(2000).		
DR	EMBL; AE002332; AAF73580.1; -		
DR	TIGR; TC0642; -		
DR	InterPro; IPR000534; Semialdh_dh.		
DR	Pfam; PF01118; Semialdehyde_dh; 1.		
DR	Pfam; PF02774; Semialdehyde_dhc; 1.		
DR	TIGRFAMS; TIGR00978; asd_EA; 1.		
KW	Complete proteome.		
SQ	SEQUENCE 334 AA; 36718 MW; 1FE91A6BFB4633E0 CRC64;		
Query Match	44.5%;	Score 40.5;	DB 16; Length 334;
Best Local Similarity	50.0%;	Pred. No. 16;	
Matches 10;	Conservative 2;	Mismatches 7;	Indels 1; Gaps
Qy	2	NTMEHVXXXESIIIXQETKYK	21
		: :	
Dd	187	NTIPHILGEER-IHQETLK	205
RESULT 9			
Q9KDHS			
ID	Q9KDHS	PRELIMINARY;	PRT; 300 AA.
AC	Q9KDHS;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Cation efflux system (zinc/cadmium).		
GN	BHI238.		
OS	Bacillus halodurans.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;		
OC	Bacillaceae; Bacillus.		
OX	NCBI_TaxID=86665;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C-125 / JCM 9153;		
RX	MEDLINE=20512582; PubMed=11058132;		
RA	Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,		
RA	Fuji F., Hirama C., Nakamura Y., Ogawara N., Kuhara S.,		
RA	Horikoshi K.;		
RT	"Complete genome sequence of the alkaliphilic bacterium Bacillus		
RT	halodurans and genomic sequence comparison with Bacillus subtilis."		
RL	Nucleic Acids Res. 28:4317-4331(2000).		
DR	EMBL; AP001511; BAB04957.1; -		
DR	InterPro; IPR002524; Cation_efflux.		


```

RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=98290299; PubMed=9628335;
RA  Lee K.F., Shaw P.C., Picone S.J., Wilson G.G., Lunnen K.D.;
RT  "Sequence comparison of the EcoHK3II and EaeI restriction-modification
RL  systems suggests an intergenic transfer of genetic material.";
DR  Biol. Chem. 379:437-441(1998).
KW  EMBL; AF039582; AAB95337.1; -.
SQ  Methyltransferase; Transferase.
    SEQUENCE 176 AA; 20348 MW; 9EDB5E3C97B6B930 CRC64;

Query Match      41.8%; Score 38; DB 2; Length 176;
Best Local Similarity 41.2%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY  1 KNTMEHVXXXEEIIXQ 17
    :||:| | | |
Db  51 QNTLEHPSVEECTLSQ 67

Search completed: February 11, 2003, 18:19:46
Job time : 25.5979 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:06:55 ; Search time 7.26804 seconds
(without alignments)
91.669 Million cell updates/sec

Title: US-09-380-738A-5
Perfect score: 13
Sequence: 1 XXXEE 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	76.9	3	AAW56195	Anti-inflammatory
2	10	76.9	3	AAW56182	Anti-inflammatory
3	10	76.9	3	AAW78378	Isoelectric focus
4	10	76.9	3	ABB47202	OB-cadherin CAR pe
5	10	76.9	4	AAP60670	Sequence of amino
6	10	76.9	4	AAP60497	Peptide with Fc re
7	10	76.9	4	AA44299	Hemoregulatory pe
8	10	76.9	4	AA452607	Inhibitor of signa
9	10	76.9	4	AAW55019	CDR3 region of L3T
10	10	76.9	4	AAW55021	CDR3 region of L3T

11	10	76.9	4	15	AAW48248	Endoplasmic reticu
12	10	76.9	4	15	AAW51489	Human p53 CK II si
13	10	76.9	4	15	AAW52963	Human p53 CK II si
14	10	76.9	4	15	AAW59898	2-amino-6,7-dihydr
15	10	76.9	4	15	AAW59899	2-amino-6,7-dihydr
16	10	76.9	4	15	AAW59889	2-amino-6,7-dihydr
17	10	76.9	4	15	AAW59890	2-amino-6,7-dihydr
18	10	76.9	4	15	AAW59892	2-amino-6,7-dihydr
19	10	76.9	4	15	AAW59893	2-amino-6,7-dihydr
20	10	76.9	4	15	AAW64036	Acylated polyhydro
21	10	76.9	4	15	AAW64042	Acylated polyhydro
22	10	76.9	4	15	AAW64043	Acylated polyhydro
23	10	76.9	4	17	AAW87075	Heterologous prote
24	10	76.9	4	18	AAW29245	High affinity Src
25	10	76.9	4	18	AAW24497	Immobilised fluore
26	10	76.9	4	19	AAW47838	Oligopeptide used
27	10	76.9	4	19	AAW78525	SH2 domain binding
28	10	76.9	4	19	AAW78527	SH2 domain binding
29	10	76.9	4	19	AAW78608	SH2 domain binding
30	10	76.9	4	19	AAW51469	Pre-sequence which
31	10	76.9	4	19	AAW51470	Pre-sequence which
32	10	76.9	4	19	AAW51471	Pre-sequence which
33	10	76.9	4	19	AAW51472	Pre-sequence which
34	10	76.9	4	19	AAW51473	Pre-sequence which
35	10	76.9	4	19	AAW51468	Pre-sequence which
36	10	76.9	4	19	AAW51463	Pre-sequence which
37	10	76.9	4	19	AAW51466	Pre-sequence which
38	10	76.9	4	19	AAW49984	Conantokin peptide
39	10	76.9	4	19	AAW38340	Peptide used in pr
40	10	76.9	4	20	AAW33530	Human apopain pept
41	10	76.9	4	20	AAW28719	Human hepatitis C
42	10	76.9	4	20	AAW28732	Human hepatitis C
43	10	76.9	4	20	AAW27155	Peptide Seq ID No:
44	10	76.9	4	20	AAW25358	Factor xa peptide
45	10	76.9	4	20	AAW15657	Peptide used to ma

ALIGNMENTS

RESULT 1
AAW56195
ID AAW56195 standard; peptide; 3 AA.
XX
AC AAW56195;
XX
DT 20-JUL-1998 (first entry)
XX
DE Anti-inflammatory tripeptide.
XX
KW Anti-inflammatory; macrophage inhibitory activity; fibronectin;
KW T-cell inhibitory activity; adherence; extracellular matrix;
KW up-regulation; fas receptor expression; inflammation.
XX
OS Synthetic.
XX
PN WO9809985-A2.
XX
PD 12-MAR-1998.
XX
PF 03-SEP-1997; 97WO-IL00295.
XX
PR 28-MAY-1997; 97US-0864301.
PR 03-SEP-1996; 96US-0025376.
PR 20-NOV-1996; 96US-0753141.
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Beserman P, Eisenbachschwartz M, Hirschberg DL;
XX
DR WPI; 1998-193550/17.
XX
PT Anti-inflammatory peptides and derivatives - used for treating, e.g.

PT arthritis, ulcerative colitis, auto-immune disease, allergy asthma,
 XX shock, HIV infection, transplant rejection or Alzheimer's disease
 XX Claim 5; Page 34; 42pp; English.
 XX
 CC AAW56171-248 represent anti-inflammatory tripeptides of the invention.
 CC They are derived from the formulae:
 CC Xaa-Glu-Arg, Arg-Glu-Xaa, Xaa-Arg-Glu, or Glu-arg-Xaa, where
 CC Xaa = any amino acid residue.
 CC Cyclic derivatives of the peptides also function as anti-inflammatory
 CC agents. The peptides can be covalently linked to one another either
 CC directly or through a spacer. The peptides and their derivatives have
 CC macrophage inhibitory and T-cell inhibitory activity and thus,
 CC anti-inflammatory activity. The peptides and compositions have
 CC anti-immune activity, i.e. inhibitory effects against a cellular and
 CC humoral immune response, including a response not associated with
 CC inflammation. The peptides also inhibit the ability of macrophages and
 CC T-cells to adhere to extracellular matrix components and fibronectin, as
 CC well as up-regulated fas receptor expression in T-cells. They can be
 CC used to inhibit unwanted immune reaction and inflammation.
 XX
 SQ Sequence 3 AA;

Query Match 76.9%; Score 10; DB 19; Length 3;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05; Length 3;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 II
 Db 2 EE 3

RESULT 2
 AAW56182
 ID AAW56182 standard; peptide; 3 AA.
 AC AAW56182;
 XX
 DT 20-JUL-1998 (first entry)
 DE
 DE Anti-inflammatory tripeptide.
 XX
 KW Anti-inflammatory; macrophage inhibitory activity; fibronectin;
 KW T-cell inhibitory activity; adherence; extracellular matrix;
 KW up-regulation; fas receptor expression; inflammation.
 XX
 OS Synthetic.
 XX
 PN WO9809985-A2.
 XX
 PD 12-MAR-1998.
 XX
 PF 03-SEP-1997; 97WO-IL00295.
 XX
 PR 28-MAY-1997; 97US-0864301.
 PR 03-SEP-1996; 96US-0025376.
 PR 20-NOV-1996; 96US-0753141.
 XX
 PA (YEDA) YEDA RES & DEV CO LTD.
 XX
 PI Beserman P, Eisenbachschwartz M, Hirschberg DL;
 XX
 DR WPI; 1998-193550/17.

XX Anti-inflammatory peptides and derivatives - used for treating, e.g.
 PT arthritis, ulcerative colitis, auto-immune disease, allergy asthma,
 PT shock, HIV infection, transplant rejection or Alzheimer's disease
 XX
 PS Claim 3; Page 34; 42pp; English.
 XX
 CC AAW56171-248 represent anti-inflammatory tripeptides of the invention.
 CC They are derived from the formulae:
 CC Xaa-Glu-Arg, Arg-Glu-Xaa, Xaa-Arg-Glu, or Glu-arg-Xaa, where

CC Xaa = any amino acid residue.
 CC Cyclic derivatives of the peptides also function as anti-inflammatory
 CC agents. The peptides can be covalently linked to one another either
 CC directly or through a spacer. The peptides and their derivatives have
 CC macrophage inhibitory and T-cell inhibitory activity and thus,
 CC anti-inflammatory activity. The peptides and compositions have
 CC anti-immune activity, i.e. inhibitory effects against a cellular and
 CC humoral immune response, including a response not associated with
 CC inflammation. The peptides also inhibit the ability of macrophages and
 CC T-cells to adhere to extracellular matrix components and fibronectin, as
 CC well as up-regulated fas receptor expression in T-cells. They can be
 CC used to inhibit unwanted immune reaction and inflammation.
 XX

SQ Sequence 3 AA;
 Query Match 76.9%; Score 10; DB 19; Length 3;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05; Length 3;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 II
 Db 1 EE 2

RESULT 3
 AAW78378
 ID AAW78378 standard; peptide; 3 AA.
 XX
 AC AAW78378;
 XX
 DT 11-MAY-1999 (first entry)
 XX
 DE
 DE Isoelectric focusing marker peptide #42.
 XX
 KW Marker peptide; capillary isoelectric focusing electrophoresis; IEF; pI;
 KW tryptophan; detection; UV absorption.
 XX
 OS Synthetic.
 XX
 PN JP11023531-A.
 XX
 PD 29-JAN-1999.
 XX
 PF 03-JUL-1997; 97JP-0178579.
 XX
 PR 03-JUL-1997; 97JP-0178579.
 XX
 PA (BUNSHI) BUNSHI BIOTONICS KENKYUSHO KK.
 XX
 DR WPI; 1999-170244/15.

XX Markers for isoelectric point electrophoresis in UV absorption
 PT detection - contain at least one tryptophan
 XX
 PS Claim 3; Page 2; 7pp; Japanese.
 XX
 CC This peptide is a synthetic marker peptide used for isoelectric focusing
 CC electrophoresis, particularly capillary isoelectric focusing
 CC electrophoresis. The peptide contains at least one tryptophan amino
 CC acid residue for detection in the UV absorption range. The markers
 CC (AAW78364-W78379) provide sharp isolation, large UV absorption intensity
 CC and wide range of pI, and can be easily handled.
 XX

SQ Sequence 3 AA;
 Query Match 76.9%; Score 10; DB 20; Length 3;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05; Length 3;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 II
 Db 2 EE 3

RESULT 4
 ABB47202
 ID ABB47202 standard; Peptide; 3 AA.
 XX
 AC ABB47202;
 XX
 DT 30-JAN-2002 (first entry)
 XX
 DE OB-cadherin CAR peptide 15.
 XX
 KW Desmosomal cadherin; cell adhesion; CAR sequence; immunosuppressive;
 KW cytostatic; antiapoptotic; wound healing; reduce scar tissue;
 KW skin graft; organ implant; autoimmune blistering disorder; cancer;
 KW apoptosis.
 XX
 OS Synthetic.
 XX
 PN WO200172956-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 27-MAR-2001; 2001WO-IB01400.
 XX
 PR 27-MAR-2000; 2000US-0535852.
 XX
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX
 PI Blaschuk OW, Symonds JM, Gour BJ;
 XX
 DR WPI; 2002-025778/03.
 XX
 PT Modulating agents for inhibiting or enhancing desmosomal cadherin
 PT mediated cell adhesion, useful for facilitating wound healing and/or
 PT reducing scar tissue, treating cancer and inducing apoptosis -
 XX
 PS Disclosure; Page 50; 127pp; English.
 XX
 CC The invention relates to modulating agents for inhibiting or enhancing
 CC desmosomal cadherin mediated cell adhesion, comprising a modulating agent
 CC comprising a desmosomal cadherin cell adhesion recognition CAR sequence
 CC (ABB45341-ABB47262), a non-peptide mimetic of a desmosomal cadherin CAR
 CC sequence, a substance such as an antibody or antigen-binding fragment
 CC that specifically binds a desmosomal cadherin CAR sequence and/or a
 CC polynucleotide encoding a polypeptide that comprises a desmosomal
 CC cadherin CAR sequence or analogue. The modulating agents have
 CC immunosuppressive, cytostatic and antiapoptotic activity and are used to
 CC facilitate wound healing and/or reduce scar tissue, for enhancing
 CC adhesion of foreign tissue implants (e.g. skin graft or organ implant),
 CC treating an autoimmune blistering disorder and to treat cancer
 CC (e.g. carcinoma, leukaemia or melanoma) and induce apoptosis.
 XX
 SQ Sequence 3 AA;
 Query Match 76.9%; Score 10; DB 23; Length 3;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 DB 1 EE 2
 RESULT 5
 AAP60670
 ID AAP60670 standard; Protein; 4 AA.
 XX
 AC AAP60670;
 XX
 DT 03-OCT-2002 (updated)
 DT 26-JUL-1991 (first entry)
 XX
 DE Sequence of amino terminal extension 3 for recombinant human growth

hormone, cleavable with dipeptidyl aminopeptidase I (DAP I).
 Recombinant human growth hormone.
 Homo sapiens.
 WO8604609-A.
 PD 14-AUG-1986.
 XX
 PF 06-FEB-1986; 86WO-DK00014.
 XX
 PR 07-FEB-1985; 85DK-0000556.
 PR 24-JUN-1988; 88DK-0003482.
 XX
 PA (NGEN-) NORDISK GENTOFTE AS.
 PA (ANDE/) ANDERSEN H D.
 PA (NOGE-) NORDISK GENTOFTE A/S.
 PA (NOVO) NOVO-NORDISK A/S.
 XX
 PI Andersen H, Pedersen J, Christensen T, Hansen J, Jessen T;
 XX
 DR WPI; 1986-225461/34.
 XX
 PT Human growth hormone prodn. - by reacting amino terminal extended
 PT hormone with dipeptidyl aminopeptidase
 XX
 PS Disclosure; Page 4; 20pp; English.
 XX
 CC The amino terminal extension is cleaved selectively and in a high
 CC yield and the formed hGH may then easily be separated from any
 CC residues of partly converted amino terminal extended hGH by anion
 CC exchange.
 CC (Updated on 03-OCT-2002 to add missing OS field.)
 XX
 SQ Sequence 4 AA;
 Query Match 76.9%; Score 10; DB 7; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 DB 3 EE 4
 RESULT 6
 AAP60497
 ID AAP60497 standard; protein; 4 AA.
 XX
 AC AAP60497;
 XX
 DT 28-JUL-1991 (first entry)
 XX
 DE Peptide with Fc receptor-blocking activity.
 XX
 KW Antiinflammatory; antiallergic; immunosuppressive.
 XX
 OS Synthetic.
 XX
 PN WO8601211-A.
 XX
 PD 27-FEB-1986.
 XX
 PF 10-AUG-1984; 84WO-EP00242.
 XX
 PR 10-AUG-1984; 84WO-EP00242.
 XX
 PA (MERE) MERCK PATENT GMBH.
 XX
 PI Hahn GS;
 XX
 DR WPI; 1986-068963/10.

XX New active site peptide derivs. - blocking binding of immune
PT complex or immunoglobulin to Fc receptors, useful e.g. for
PT treating auto-immune disease.

XX Claim 22; Page 78-79; 87pp; English.

XX The sequence is an active-site compound which blocks immune-
CC complex binding to Ig Fc receptors and/or Ig binding to lymphocyte
CC Fc receptors. The peptide modulates immune complex-mediated
CC immunosuppression, inflammation and tissue disruption, and reduces
CC human allergic responses.

XX Sequence 4 AA;

Query Match 76.9%; Score 10; DB 7; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
Db 3 EE 4

RESULT 7

AAR44299
ID AAR44299 standard; peptide; 4 AA.

XX AAR44299;

XX 13-JUN-1994 (first entry)

XX Hemoregulatory peptide.

XX Hemoregulatory peptide; inhibition; cell division;
KW myelopoietic cells; bone marrow cells; proliferation; cancer;
KW autoimmune disease; psoriasis; antibody; immunoassay.

OS Synthetic.

XX Key Location/Qualifiers
FT Modified-site 3 /note= "Asp-NH-CH(-A)-CO-, A is a group -C-RA-RA'-Z,
FT RA = H or a group RA', -O-RA", -S-RA",
FT -N-RA'-RA", -C-O-N-RA"-RA" or -C-O-O-RA";
FT RA' = H or a group RA"; RA" = an alkyl,
FT cycloalkyl, alkanoyl, hydroxyalkyl, anidine
FT group, or a carboxylic or heterocyclic group;
FT Z = a group -O-RB, -N-RC-RC, C-RD-RE-RF, or a
FT substituted or unsubstituted, aromatic or non-
FT aromatic, carbocyclic or heterocyclic ring
FT other than unsubstituted phenyl; RB = a
FT unsaturated hydrocarbon group such as an alkyl,
FT aralkyl or aryl group, optionally substituted
FT by one or more RA groups, and optionally
FT interrupted by one or more -N-, -O- OR -S-
FT heteroatoms; RC= H or a group RB; RD= H or a
FT group RF; RE= H or a group RF or together with
FT group RD forms a C=O group; RF = a group -RB,
FT -O-RB, -N-RC-RC or -S-RB, or -OH, carboxy,
FT aminocarbonyl or alkoxy group, or is a
FT methylene group linked to a N atom attached to
FT the alpha atom, or together with RD forms an
FT alkylidene group, or is H where one of RA, RA',
FT RD and RE is other than hydrogen"

XX WO9324524-A.

XX 09-DEC-1993.

XX 02-JUN-1993; 93WO-GB01172.

PT . . .

PR 02-JUN-1992; 92GB-0011668.

XX (HAFS-) HAFSLUND NYCOMED AS.

PA (HOLM/) HOLMES M J.

XX Undheim K;

DR WPI; 1993-405729/50.

XX Single-chain haemo-regulatory peptide derivs. - useful for
PT inhibiting proliferation of cells, partic. myelopoietic and bone
PT marrow cells

XX Claim 11; Page 46; 54pp; English.

XX This sequence represents a hemoregulatory peptide compound
CC derivative. Peptides based on this sequence may be used for
CC inhibiting cell division, particularly myelopoietic or bone marrow
CC cells. They can be used to inhibit proliferation of cells in the
CC treatment of cancer, autoimmune diseases or psoriasis. They may
CC also be used to generate antibodies for use in immunoassays.

XX Sequence 4 AA;

Query Match 76.9%; Score 10; DB 14; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5

Db 1 EE 2

RESULT 8

AAR52607

ID AAR52607 standard; peptide; 4 AA.

XX AAR52607;

XX 23-DEC-1994 (first entry)

XX Inhibitor of signal transduction.

XX Signal transduction; phosphonomethylphenylalanine;
KW monofluorophosphonomethylphenylalanine;
KW difluorophosphonomethylphenylalanine;
KW protein tyrosine phosphatase; kinase; enzyme;
KW hydrolysis resistant; phosphorous; diabetes; cancer;
KW malignancy.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "phosphonomethylphenylalanine
FT or mono- or difluorophosphono-
FT methylphenylalanine"

XX WO9408600-A.

XX 28-APR-1994.

XX 08-OCT-1993; 93WO-US09626.

XX 09-OCT-1992; 92US-0959949.

XX (JOSL-) JOSLIN DIABETES CENT.

XX Shoelson S;

XX WPI; 1994-150930/18.

PT New peptide(s) contg. tyrosine or phosphorylated analogues - used

PT for inhibiting the dephosphorylation of a substrate by a protein
 XX tyrosine phosphatase
 PS Claim 12; Page 57; 70pp; English.

XX A peptide capable of inhibiting the interaction of a SH2-domain
 CC contg. protein with a second protein comprises sequence (I)
 CC R1-R2-R3-R4 (I)
 CC R1= Tyr, phosphotyrosine or an analogue of phosphotyrosine
 CC having a hydrolysis resistant phosphorous moiety;
 CC R2= Asp, Thr, Tyr, His, Gln, Met, Val, Ile or Glu;
 CC R3= any amino acid;
 CC R4= Ile, Met, Leu or Val.
 CC Examples of such peptides are given in AAR52607-19.
 CC The peptides can be used for inhibiting cellular protein
 CC tyrosine phosphatases (PTases) and for controlling metabolic
 CC processes, e.g. abnormal processes associated with diabetes and
 CC for treating selected malignancies. They can also be used to
 CC study the enzymatic mechanisms of PTase activity and to
 CC investigate the metabolic and biochemical roles of PTases.
 XX SQ

Sequence 4 AA;

Query Match 76.9%; Score 10; DB 15; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 9

AAR55019 ID AAR55019 standard; peptide; 4 AA.

XX AC AAR55019;

XX DT 16-DEC-1994 (first entry)

XX DE CDR3 region of L3T4 peptide mimic.

XX CD4; Complementarity determining regions; CDR; lateral domains;
 KW mimics; glycoproteins; immunoglobulin superfamily; SLE; MS; RA; GVH;
 KW inhibit T cell proliferation; treatment of multiple sclerosis;
 KW systemic lupus erythematosus; graft rejection; rheumatoid arthritis;
 KW graft versus host disease; T cell leukaemias.

XX OS Synthetic.

XX PN WO9411014-A.

XX PD 26-MAY-1994.

XX PF 12-NOV-1993; 93WO-US10999.

XX PR 13-NOV-1992; 92US-0977692.

XX PR 11-JUN-1993; 93US-0076092.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Jameson BA, Korngold R, McDonnell JM;

XX WPI; 1994-183151/22.

XX Compounds displaying a surface similar to the surface presented
 PT by one of five distinct lateral domains of CD4 - inhibit T cell
 PT proliferation

PS Claim 3; Page 50; 78pp; English.

XX A molecular model of the mouse CD4 protein (L3T4) was developed
 CC from the high resolution crystal structure of human CD4 (Brookhaven

CC code: 1CD4). The CDR regions were used for modelling peptide mimics.
 CC CDR3 peptides comprise the essential amino acids 88-90 (Asp Gln Lys)
 CC or 90-92 (Lys Glu Glu) of CD4 and may comprise additional flanking
 CC sequences from CD4. This peptide comprises amino acids 89-92.
 CC The peptide mimics are useful for inhibiting the proliferation of
 CC T cells modulating immune responses in mammals and may be used to
 CC treat, e.g. SLE, RA, MS, GVH, graft rejection and T cell leukaemias,
 CC etc.

XX SQ Sequence 4 AA;

Query Match 76.9%; Score 10; DB 15; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 3 EE 4

RESULT 10

AAR55021 ID AAR55021 standard; peptide; 4 AA.

XX AC AAR55021;

XX DT 16-DEC-1994 (first entry)

XX DE CDR3 region of L3T4 peptide mimic.

XX CD4; Complementarity determining regions; CDR; lateral domains;
 KW mimics; glycoproteins; immunoglobulin superfamily; SLE; MS; RA; GVH;
 KW inhibit T cell proliferation; treatment of multiple sclerosis;
 KW systemic lupus erythematosus; graft rejection; rheumatoid arthritis;
 KW graft versus host disease; T cell leukaemias.

XX OS Synthetic.

XX PN WO9411014-A.

XX PD 26-MAY-1994.

XX PF 12-NOV-1993; 93WO-US10999.

XX PR 13-NOV-1992; 92US-0977692.

XX PR 11-JUN-1993; 93US-0076092.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Jameson BA, Korngold R, McDonnell JM;

XX WPI; 1994-183151/22.

XX Compounds displaying a surface similar to the surface presented
 PT by one of five distinct lateral domains of CD4 - inhibit T cell
 PT proliferation

PS Claim 3; Page 50; 78pp; English.

XX A molecular model of the mouse CD4 protein (L3T4) was developed
 CC from the high resolution crystal structure of human CD4 (Brookhaven
 CC code: 1CD4). The CDR regions were used for modelling peptide mimics.
 CC CDR3 peptides comprise the essential amino acids 88-90 (Asp Gln Lys)
 CC or 90-92 (Lys Glu Glu) of CD4 and may comprise additional flanking
 CC sequences from CD4. This peptide comprises amino acids 90-93.
 CC The peptide mimics are useful for inhibiting the proliferation of
 CC T cells modulating immune responses in mammals and may be used to
 CC treat, e.g. SLE, RA, MS, GVH, graft rejection and T cell leukaemias,
 CC etc.

XX SQ Sequence 4 AA;

Query Match 76.9%; Score 10; DB 15; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EE 5
II
Db 2 EE 3

RESULT 11

AAR48248
ID AAR48248 standard; peptide; 4 AA.

XX
AC AAR48248;
XX

DT 29-JUL-1994 (first entry)

XX Endoplasmic reticulum retention signal DEEL.

DE Single chain antibody; sFv; heavy chain; light chain; kappa;
KW variable domain; hydrophilic linker; antibodies;
KW endoplasmic reticulum retention.

XX Synthetic.

XX WO9402610-A.

XX 03-FEB-1994.

XX 16-JUL-1993; 93WO-US06735.

XX 17-JUL-1992; 92US-0916939.

XX 17-MAR-1993; 93US-0045274.

XX (DAND) DANA FARBER CANCER INST INC.

XX Haseltine WA, Marasco WA;

XX WPI; 1994-048868/06.

XX Intracellular binding of antigens - by using antibody targeting
PT with vector system, for e.g. tumour suppression

XX Claim 32; Page 100; 155pp; English.

XX New vector systems comprise a sequence adapted for intracellular
CC delivery and expression contg. a promoter operably linked to an
CC antibody gene encoding an antibody which binds to a specific target
CC antigen. The antibody is esp. a single chain antibody in which the
CC heavy and light chain variable regions are joined via a hydrophilic
CC linker peptide. Localisation sequences are pref. included in the
CC constructs. See AAR48246-9 and AAR48252-3 for pref. (known) endoplasmic
XX reticulum retention signals.

SQ Sequence 4 AA;

Query Match 76.9%; Score 10; DB 15; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EE 5
II
Db 2 EE 3

RESULT 12

AAR51489
ID AAR51489 standard; Peptide; 4 AA.

XX
AC AAR51489;

XX 10-SEP-1994 (first entry)

XX Human p53 CK II site.

XX

XX

KW p53; CK II; casein-kinase II; phosphorylation; CcN motif;
KW TLE; transducin-like enhancer of split protein;
KW protein transport; differentiation; cervix cancer; dysplasia;
KW malignancy.

XX Homo sapiens.

XX WO9408037-A.

XX 14-APR-1994.

XX 30-SEP-1993; 93WO-US09333.

XX 30-SEP-1992; 92US-0955011.

XX (MEDI-) MEDICAL RES COUNCIL.
XX (UYA) UNIV YALE.

XX Artavanis-Issakonas S, Hill RE, Redhead NJ, Stifani S;

XX WPI; 1994-135597/16.

XX New human transducin-like enhancers of split protein - and
PT associated multi-protein complexes, chimeric proteins,
PT antibodies, nucleic acid, etc., involved in nuclear-cytoplasmic
PT protein transport

XX Disclosure; Page 35; 112pp; English.

XX The nucleotide and deduced aa sequences of human transducin-like
CC enhancer of split proteins TLE-1 (AAQ62175, AAR51476), TLE-2 (AAQ62176,
CC AAR51477), TLE-3 (AAQ62177, AAR51478) and TLE-4 (AAQ62178, AAR51479)
CC were determined. The aa sequences were compared with that of Drosophila
CC E(spl)m9/10 protein (AAR51481). Comparison of the WD-40 domains of
CC these proteins defined the consensus residues shown in AAR51480. The
CC CcN motifs of the proteins were compared with those of the SV40 T
CC antigen, human c-myc, human p53, human A-myb and dorsal proteins
CC with respect to nuclear localization site, and casein-kinase II and
CC cdc2-kinase phosphorylation sites (sequences AAR51482-96). TLE can be
CC used to treat or diagnose (pre)neoplastic conditions, or
CC to study cell differentiation.

SQ Sequence 4 AA;

Query Match 76.9%; Score 10; DB 15; Length 4;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EE 5
II
Db 2 EE 3

RESULT 13

AAR52963
ID AAR52963 standard; Peptide; 4 AA.

XX
AC AAR52963;

XX 10-SEP-1994 (first entry)

XX Human p53 CK II site.

XX p53; CK II; casein-kinase II; phosphorylation; CcN motif;
KW transducin-like enhancer of split protein; TLE; cell fate;
KW differentiation; cervix cancer; breast cancer; psoriasis; baldness.

XX Homo sapiens.

XX WO9407522-A.

XX 14-APR-1994.

XX 30-SEP-1993; 93WO-US09339.
 XX 30-SEP-1992; 92US-0954813.
 XX (UYVA) UNIV YALE.
 XX Artavanis-tsakonas S, Stifani S;
 XX WPI; 1994-135221/16.
 XX Transducin-like enhancer or split proteins and nucleic acids -
 XX are for treatment of disorders of cell fate or differentiation
 XX e.g. cervical cancer, breast cancer, psoriasis, baldness etc.
 XX Disclosure; Page 62; 147pp; English.
 XX The nucleotide and deduced aa sequences of human transducin-like
 XX enhancer of split proteins TLE-1 (AAQ45333, AAR51109), TLE-2 (AAQ45334,
 XX AAR51110), TLE-3 (AAQ45335, AAR51111) and TLE-4 (AAQ45336, AAR52953)
 XX were determined. The aa sequences were compared with that of
 XX Drosophila E(spl) m9/10 (AAR52955). Comparison of the WD-40 domains
 XX of these proteins defined the consensus residues shown in AAR52954.
 XX The CcN motifs of the proteins were compared with those of SV40 T
 XX antigen, human c-myc, human p53, human A-myb and dorsal protein with
 XX respect to nuclear localization site, and casein-kinase and cdk2-
 XX kinase phosphorylation sites (sequences AAR52956-70).
 XX Sequence 4 AA;
 Query Match 76.9%; Score 10; DB 15; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 DB 2 EE 3
 RESULT 14
 AAR59898
 ID AAR59898 standard; peptide; 4 AA.
 AC AAR59898;
 XX 17-FEB-1995 (first entry)
 DT 2-amino-6,7-dihydroxy-4-thiaheptanoic acid deriv.34.
 DE immunostimulant; thrombocytopaenia; haematopoiesis insufficiency;
 XX 2-amino-6,7-dihydroxy-4-thiaheptanoic acid; bone marrow;
 KW transplantation.
 KW Synthetic.
 OS Key Location/Qualifiers
 XX Modified-site 1
 FT /label= OTHER
 FT /note= "(2R,6R)-2-amino-6,7-bis(Steo)-4-
 FT thiaheptanoyl",
 FT where Steo = octadecanoyloxy"
 XX EP604957-A.
 PN 06-JUL-1994.
 PD 27-DEC-1993; 93EP-0120970.
 PF 28-DEC-1992; 92JP-0349062.
 XX 16-MAR-1993; 93JP-0056185.
 XX 22-JUL-1993; 93JP-0181735.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Harada S, Hida T, Tanida S, Wakimasu M, Yukishige K;
 XX WPI; 1994-210219/26.
 XX New 2-amino-6,7-di:hydroxy-4-thia-heptanoic acid derivs. - are

XX Harada S, Hida T, Tanida S, Wakimasu M, Yukishige K;
 XX WPI; 1994-210219/26.
 XX New 2-amino-6,7-di:hydroxy-4-thia-heptanoic acid derivs. - are
 XX immuno-stimulating agents useful for treating e.g.
 XX thrombocytopenia
 XX Example 42; Page 81; 89pp; English.
 XX This peptide is a specific example of a generic formula covering
 XX 2-amino-6,7-dihydroxy-4-thiaheptanoic acid derivs. bonded to a
 XX sequence of 1-10 amino acid residues, at least one of which has a
 XX water-solubility enhancing group. Such compounds improve the state
 XX of haematopoiesis-insufficiency and can be used for treating or
 XX preventing leukocytopenia caused by radiotherapy or chemotherapy
 XX of cancers. They can also be used as haematopoietic stimulating
 XX agents in the case of bone marrow transplantation, as immuno-
 XX stimulating agents having leukocyte-increasing action; and for
 XX treating thrombocytopenia.
 XX Sequence 4 AA;
 Query Match 76.9%; Score 10; DB 15; Length 4;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 DB 3 EE 4
 RESULT 15
 AAR59899
 ID AAR59899 standard; peptide; 4 AA.
 AC AAR59899;
 XX 17-FEB-1995 (first entry)
 DT 2-amino-6,7-dihydroxy-4-thiaheptanoic acid deriv.35.
 DE immunostimulant; thrombocytopaenia; haematopoiesis insufficiency;
 XX 2-amino-6,7-dihydroxy-4-thiaheptanoic acid; bone marrow;
 KW transplantation.
 KW Synthetic.
 OS Key Location/Qualifiers
 XX Modified-site 1
 FT /label= OTHER
 FT /note= "(2R,6R)-2-amino-6,7-bis(Myro)-4-
 FT thiaheptanoyl",
 FT where Myro = tetradecanoyloxy"
 XX EP604957-A.
 PN 06-JUL-1994.
 PD 27-DEC-1993; 93EP-0120970.
 PF 28-DEC-1992; 92JP-0349062.
 XX 16-MAR-1993; 93JP-0056185.
 XX 22-JUL-1993; 93JP-0181735.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Harada S, Hida T, Tanida S, Wakimasu M, Yukishige K;
 XX WPI; 1994-210219/26.
 XX New 2-amino-6,7-di:hydroxy-4-thia-heptanoic acid derivs. - are

PT immuno-stimulating agents useful for treating e.g.
PT thrombocytopenia

XX

PS Example 43; Page 81; 89pp; English.

XX

CC This peptide is a specific example of a generic formula covering
CC 2-amino-6,7-dihydroxy-4-thiaheptanoic acid derivs. bonded to a
CC sequence of 1-10 amino acid residues, at least one of which has a
CC water-solubility enhancing group. Such compounds improve the state
CC of haematopoiesis-insufficiency and can be used for treating or
CC preventing leukocytopenia caused by radiotherapy or chemotherapy
CC of cancers. They can also be used as haematopoietic stimulating
CC agents in the case of bone marrow transplantation, as immuno-
CC stimulating agents having leukocyte-increasing action; and for
CC treating thrombocytopenia.

XX

SQ Sequence 4 AA;

Query Match 76.9%; Score 10; DB 15; Length 4;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5

Db 3 EE 4

Search completed: February 11, 2003, 18:16:57

Job time : 8.26804 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:14:31 ; Search time 2.47423 Seconds
(without alignments)
59.459 Million cell updates/sec

Title: US-09-380-738A-5
Perfect score: 13
Sequence: 1 XXXEE 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	76.9	3	1	US-08-126-564A-21
2	10	76.9	3	1	US-08-174-365A-59
3	10	76.9	3	2	US-08-539-432-8
4	10	76.9	3	5	PCT-US94-09143-21
5	10	76.9	4	1	US-07-906-349A-15
6	10	76.9	4	1	US-08-174-365A-28
7	10	76.9	4	1	US-08-174-365A-61
8	10	76.9	4	1	US-08-174-365A-62
9	10	76.9	4	1	US-08-174-365A-67
10	10	76.9	4	1	US-08-174-365A-103
11	10	76.9	4	1	US-08-174-365A-105
12	10	76.9	4	1	US-08-178-570-43
13	10	76.9	4	1	US-07-807-529A-35
14	10	76.9	4	1	US-07-872-673B-11
15	10	76.9	4	1	US-08-076-092-27
16	10	76.9	4	1	US-08-076-092-29
17	10	76.9	4	1	US-07-969-305-57
18	10	76.9	4	1	US-07-969-305-58
19	10	76.9	4	1	US-07-969-305-59
20	10	76.9	4	1	US-07-969-305-61
21	10	76.9	4	1	US-07-969-305-62
22	10	76.9	4	1	US-07-969-305-63
23	10	76.9	4	1	US-07-969-305-68
24	10	76.9	4	1	US-07-969-305-69
25	10	76.9	4	1	US-07-969-305-70
26	10	76.9	4	1	US-07-789-184-82
27	10	76.9	4	1	US-07-789-184-89

28	10	76.9	4	1	US-07-789-184-93
29	10	76.9	4	1	US-08-323-170B-3
30	10	76.9	4	1	US-08-406-192-25
31	10	76.9	4	1	US-08-475-263-82
32	10	76.9	4	1	US-08-475-263-89
33	10	76.9	4	1	US-08-475-263-93
34	10	76.9	4	1	US-08-485-886-82
35	10	76.9	4	1	US-08-485-886-89
36	10	76.9	4	1	US-08-485-886-93
37	10	76.9	4	1	US-08-408-604A-3
38	10	76.9	4	1	US-08-408-604A-7
39	10	76.9	4	1	US-08-408-604A-96
40	10	76.9	4	1	US-08-456-424-49
41	10	76.9	4	2	US-08-545-151-25
42	10	76.9	4	2	US-08-477-362-82
43	10	76.9	4	2	US-08-477-362-89
44	10	76.9	4	2	US-08-477-362-93
45	10	76.9	4	2	US-08-373-190-19

ALIGNMENTS

RESULT 1
US-08-126-564A-21
; Sequence 21, Application US/08126564A
; Patent No. 5436150
; GENERAL INFORMATION:
; APPLICANT: Chandrasegaran, Srinivasan
; TITLE OF INVENTION: Functional Domains in Foki
; TITLE OF INVENTION: Restriction Endonuclease
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman, Darby & Cushman
; STREET: 1100 New York Ave., N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; SOFTWARE: Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/126,564A
; FILING DATE: 27-SEPTEMBER-93
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kokulis, Paul N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: PNK/4130/82506/CLB
; TELEPHONE: 202-861-3503
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-126-564A-21

Query Match 76.9%; Score 10; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 EE 5
Db 1 EE 2

RESULT 2

US-08-174-365A-59
 ; Sequence 59, Application US/08174365A
 ; Patent No. 5478809
 ; GENERAL INFORMATION:
 ; APPLICANT: Seiichi TANIDA et al.
 ; TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES
 ; NUMBER OF SEQUENCES: 106
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wenderoth, Lind & Ponack
 ; STREET: 805 Fifteenth Street, N.W., #700
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: Wordperfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/174,365A
 ; FILING DATE: December 28, 1993
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Warren M. Cheek, Jr.
 ; REGISTRATION NUMBER: 33,367
 ; REFERENCE/DOCKET NUMBER:
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-371-8850
 ; TELEFAX:
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 59:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: modified site
 ; LOCATION:
 ; IDENTIFICATION METHOD:
 ; OTHER INFORMATION: /note = "Xaa is modified amino acid as
 ; described in specification"
 ; US-08-174-365A-59

Query Match 76.9%; Score 10; DB 1; Length 3;
 Best Local Similarity 100.0%; Pred. No. 2e+05; 0;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EE 5
 ||
 Db 2 EE 3

RESULT 3

US-08-539-432-8
 ; Sequence 8, Application US/08539432
 ; Patent No. 5872210
 ; GENERAL INFORMATION:
 ; APPLICANT: MEDABALIMI, JOHN L.
 ; TITLE OF INVENTION: TRANSFRAME INHIBITORY
 ; ELEMENT OF VIRAL
 ; TITLE OF INVENTION: PROTEASE
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
 ; STREET: 345 PARK AVENUE

CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10054
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/539,432
 ; FILING DATE: 05-OCT-1995
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: DOROTHY R. AUTH
 ; REGISTRATION NUMBER: 36,434
 ; REFERENCE/DOCKET NUMBER: 2026-4194
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 758-4800
 ; TELEFAX: (212) 751-6849
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3
 ; TYPE: Amino Acid
 ; STRANDEDNESS: Unknown
 ; TOPOLOGY: Linear
 ; US-08-539-432-8

Query Match 76.9%; Score 10; DB 2; Length 3;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EE 5
 ||
 Db 1 EE 2

RESULT 4

PCT-US94-09143-21
 ; Sequence 21, Application PC/TUS9409143
 ; GENERAL INFORMATION:
 ; APPLICANT: Chandrasegaran, Srinivasan
 ; TITLE OF INVENTION: Functional Domains in FokI
 ; RESTRICTION ENDONUCLEASE
 ; NUMBER OF SEQUENCES: 48
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cushman, Darby & Cushman
 ; STREET: 1100 New York Ave., N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3918
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0,
 ; Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/09143
 ; FILING DATE: 23-AUG-1994
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/126,564
 ; FILING DATE: 27-SEPTEMBER-93
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kokulis, Paul N.
 ; REGISTRATION NUMBER: 16,773
 ; REFERENCE/DOCKET NUMBER: PNK/4130/82506/CLB
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-861-3503
 ; TELEFAX: 202-822-0944
 ; TELEX: 6714627 CUSH

; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 PCT-US94-09143-21

Query Match 76.9%; Score 10; DB 5; Length 3;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 Db 1 EE 2

RESULT 5
 US-07-906-349A-15
 ; Sequence 15, Application US/07906349A
 ; Patent No. 5434064
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlössinger, Joseph
 ; APPLICANT: Skolnik, Edward Y.
 ; APPLICANT: Margolis, Benjamin L.
 ; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
 ; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AN
 ; TITLE OF INVENTION: TARGET PROTEINS
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Browdy and Neimark
 ; STREET: 419 Seventh Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20004

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/906.349A
 ; FILING DATE: 30-JUN-1992
 ; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/643,237
 ; FILING DATE: 18-JAN-1991
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-628-5197
 ; TELEFAX: 202-737-3528
 ; INFORMATION FOR SEQ ID NO: 15:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-07-906-349A-15

Query Match 76.9%; Score 10; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 Db 2 EE 3

RESULT 6
 US-08-174-365A-28
 ; Sequence 28, Application US/08174365A

; Patent No. 5478809
 ; GENERAL INFORMATION:
 ; APPLICANT: Seilichi TANIDA et al.
 ; TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES
 ; NUMBER OF SEQUENCES: 106
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wenderoth, Lind & Ponack
 ; STREET: 805 Fifteenth Street, N.W., #700
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: Wordperfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/174.365A
 ; FILING DATE: December 28, 1993
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:

; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Warren M. Cheek, Jr.
 ; REGISTRATION NUMBER: 33,367
 ; REFERENCE/DOCKET NUMBER:
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-371-8850
 ; TELEFAX:

; INFORMATION FOR SEQ ID NO: 28:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

; NAME/KEY: modified site
 ; LOCATION:
 ; IDENTIFICATION METHOD:
 ; OTHER INFORMATION: /note = "Xaa is modified amino acid as
 ; OTHER INFORMATION: described in specification"
 US-08-174-365A-28

Query Match 76.9%; Score 10; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 Db 3 EE 4

RESULT 7
 US-08-174-365A-61
 ; Sequence 61, Application US/08174365A
 ; Patent No. 5478809
 ; GENERAL INFORMATION:
 ; APPLICANT: Seilichi TANIDA et al.
 ; TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES
 ; NUMBER OF SEQUENCES: 106
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Wenderoth, Lind & Ponack
 ; STREET: 805 Fifteenth Street, N.W., #700
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
 ; COMPUTER: IBM Compatible

```

; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/174,365A
; FILING DATE: December 28, 1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: modified site
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note = "xaa is modified amino acid as
; described in specification"
; US-08-174-365A-61

```

```

Query Match 76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
   ||
DB 2 EE 3

RESULT 8
US-08-174-365A-62
; Sequence 62, Application US/08174365A
; Patent No. 5478809
; GENERAL INFORMATION:
; APPLICANT: Seiichi TANIDA et al.
; TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/174,365A
; FILING DATE: December 28, 1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: modified site
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note = "xaa is modified amino acid as
; described in specification"
; US-08-174-365A-61

```

```

; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: modified site
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note = "xaa is modified amino acid as
; described in specification"
; US-08-174-365A-62

```

```

Query Match 76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
   ||
DB 2 EE 3

```

```

RESULT 9
US-08-174-365A-67
; Sequence 67, Application US/08174365A
; Patent No. 5478809
; GENERAL INFORMATION:
; APPLICANT: Seiichi TANIDA et al.
; TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/174,365A
; FILING DATE: December 28, 1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: modified site
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: /note = "xaa is modified amino acid as
; described in specification"
; US-08-174-365A-67

```

Query Match 76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 3 EE 4

RESULT 10
US-08-174-365A-103
; Sequence 103, Application US/08174365A
; Patent No. 5478809
; GENERAL INFORMATION:
; APPLICANT: Seiichi TANIDA et al.
; TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/174.365A
FILING DATE: December 28, 1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:

TELEX:
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: modified site
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: /note = "Xaa is modified amino acid as
described in specification"
US-08-174-365A-103

Query Match 76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 3 EE 4

RESULT 11
US-08-174-365A-105
; Sequence 105, Application US/08174365A
; Patent No. 5478809
; GENERAL INFORMATION:

APPLICANT: Seiichi TANIDA et al.
TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/174.365A
FILING DATE: December 28, 1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:

TELEX:
INFORMATION FOR SEQ ID NO: 105:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: modified site
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: /note = "Xaa is modified amino acid as
described in specification"
US-08-174-365A-105

Query Match 76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 2 EE 3

RESULT 12
US-08-178-570-43
; Sequence 43, Application US/08178570
; Patent No. 5532167
; GENERAL INFORMATION:
; APPLICANT: Lewis C. Cantley
; APPLICANT: Zhou Song yang
; TITLE OF INVENTION: Substrate Specificity of Protein Kinases
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109-1875

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

: SOFTWARE: ASCII text
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/178,570
: FILING DATE: JANUARY 7, 1994
: CLASSIFICATION: 435
:
: ATTORNEY/AGENT INFORMATION:
:
: NAME: DeConti, Giulio A., Jr.
: REGISTRATION NUMBER: 31,503
: REFERENCE/DOCKET NUMBER: BBI-004
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
:
: INFORMATION FOR SEQ ID NO: 43:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FRAGMENT TYPE: internal
: FEATURE:
:
: NAME/KEY: Modified-site
: LOCATION: 1
: OTHER INFORMATION: /note= "Xaa is
US-08-178-570-43

```

Query Match 76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels

Qy	4	EE	5
Db	2	EE	3

RESULT 13

```

US-07-807-529A-35
: Sequence 35, Application US/07807529A
: Patent No. 5547669
:
: GENERAL INFORMATION:
: APPLICANT: Rogers, Bruce L.
: APPLICANT: Morgenstern, Jay
: APPLICANT: Bond, Julian F.
: APPLICANT: Garman, Richard D.
: APPLICANT: Greenstein, Julia L.
: APPLICANT: Kuo, Mei-chang
: APPLICANT: Morville, Malcolm
:
: TITLE OF INVENTION: RECOMBITOPE PEPTIDES
: NUMBER OF SEQUENCES: 76
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
: STREET: One Kendall Square, Building 600
: CITY: Cambridge
: STATE: MA
: COUNTRY: USA
: ZIP: 02139
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII TEXT
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/807,529A
: FILING DATE: 199111213
: CLASSIFICATION: 514
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/662,276
: FILING DATE: 28-FEB-1991
: APPLICATION NUMBER: US 07/431,565
: FILING DATE: 03-NOV-1989
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Channing, Stacey L.
: REGISTRATION NUMBER: 31,095
: REFERENCE/DOCKET NUMBER: IPC-027/Imi-015

```

```

;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: (617) 494-0060
;
; INFORMATION FOR SEQ ID NO: 35:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 4 amino acids
;
; TYPE: AMINO ACID
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
US-07-807-529A-35

```

Query Match 76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels

QY	4	EE	5
Db	1	EE	2

RESULT 14

```

US-07-872-673B-11
; Sequence 11, Application US/07872673B
; Patent No. 5578466
;
; GENERAL INFORMATION:
;
; APPLICANT: TOSHIBA HAYANO, Setsuko KATO, No. 5578466uhiro TAKAHARA
;
; TITLE OF INVENTION: Co-expression System of Protein Disulfide Iso
;
; NUMBER OF SEQUENCES: 24
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Type 2DD, 3.50 inch, 720 KB
;
; COMPUTER: Apple Macintosh SE
; OPERATING SYSTEM: Apple DOS
; SOFTWARE: Microsoft Word Version 4.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,673B
; FILING DATE: 19920417
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: Japanese Patent Application No. 5578466. 1
; FILING DATE: 18-APR-1991 and 30-OCT-1991
;
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:

```

Query Match 76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels

QY 4 EE 5
||
Db 2 EE 3

RESULT 15

US-08-076-092-27
; Sequence 27, Application US/08076092
; Patent No. 5589458
; GENERAL INFORMATION:
; APPLICANT: Jameson, Bradford A.
; APPLICANT: McDonnell, James M.
; APPLICANT: Korngold, Robert
; TITLE OF INVENTION: Compounds That Inhibit T Cell
; TITLE OF INVENTION: Proliferation And Methods Using The Same
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5589458rls
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb Storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/076,092
; FILING DATE: 19930611
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/977,692
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-904
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4
; TYPE: Amino Acid
; TOPOLOGY: linear
US-08-076-092-27

Query Match 76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 2e+05; 0; Indels 0; Caps 0;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 4 EE 5
||
Db 3 EE 4

Search completed: February 11, 2003, 18:22:10
Job time : 2.67423 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:19:51 ; Search time 2.42268 Seconds
(without alignments)
52.729 Million cell updates/sec

Title: US-09-380-738A-5
Perfect score: 13
Sequence: 1 XXXEE 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	10	76.9	4	9	US-09-742-096-17	Sequence 17, Appl
2	10	76.9	4	9	US-09-742-096-18	Sequence 18, Appl
3	10	76.9	4	9	US-09-742-096-19	Sequence 19, Appl
4	10	76.9	4	9	US-09-178-286-21	Sequence 21, Appl
5	10	76.9	4	9	US-09-264-516A-51	Sequence 51, Appl
6	10	76.9	4	9	US-09-264-516A-52	Sequence 52, Appl
7	10	76.9	4	9	US-09-880-132-56	Sequence 56, Appl
8	10	76.9	4	9	US-10-061-395-37	Sequence 37, Appl
9	10	76.9	4	9	US-10-117-641-37	Sequence 37, Appl
10	10	76.9	4	9	US-10-087-905-5	Sequence 5, Appli
11	10	76.9	4	9	US-10-059-720-62	Sequence 62, Appl
12	10	76.9	4	9	US-10-235-552-3	Sequence 3, Appli
13	10	76.9	4	10	US-09-736-611-15	Sequence 15, Appl
14	10	76.9	4	10	US-09-740-359-1	Sequence 1, Appli
15	10	76.9	4	10	US-09-866-824A-1	Sequence 1, Appli
16	10	76.9	4	10	US-09-866-824A-4	Sequence 4, Appli
17	10	76.9	4	10	US-09-866-824A-5	Sequence 5, Appli
18	10	76.9	4	10	US-09-866-824A-6	Sequence 6, Appli
19	10	76.9	4	10	US-09-834-765-727	Sequence 727, App

20 10 76.9 4 10 US-09-875-519A-31 Sequence 31, Appl
21 10 76.9 4 10 US-09-972-599A-57 Sequence 57, Appl
22 10 76.9 4 10 US-09-234-395-25 Sequence 25, Appl
23 10 76.9 4 10 US-09-234-395-26 Sequence 26, Appl
24 10 76.9 4 10 US-09-380-323-6 Sequence 6, Appli
25 10 76.9 4 10 US-09-380-323-7 Sequence 7, Appli
26 10 76.9 4 10 US-09-894-711-1 Sequence 1, Appli
27 10 76.9 4 10 US-09-305-928-25 Sequence 25, Appl
28 10 76.9 4 10 US-09-305-928-26 Sequence 26, Appl
29 10 76.9 4 10 US-09-880-149-56 Sequence 56, Appl
30 10 76.9 4 10 US-09-947-387-16 Sequence 16, Appl
31 10 76.9 4 10 US-09-947-387-40 Sequence 40, Appl
32 10 76.9 4 12 US-10-043-142-8 Sequence 8, Appli
33 10 76.9 5 9 US-10-020-436A-4 Sequence 4, Appli
34 10 76.9 5 9 US-10-020-436A-5 Sequence 5, Appli
35 10 76.9 5 9 US-10-020-436A-8 Sequence 8, Appli
36 10 76.9 5 9 US-10-020-436A-9 Sequence 9, Appli
37 10 76.9 5 9 US-10-020-436A-11 Sequence 11, Appl
38 10 76.9 5 9 US-10-020-436A-13 Sequence 13, Appl
39 10 76.9 5 9 US-10-020-436A-14 Sequence 14, Appl
40 10 76.9 5 9 US-09-976-736-56 Sequence 56, Appl
41 10 76.9 5 9 US-09-976-736-60 Sequence 60, Appl
42 10 76.9 5 9 US-09-976-736-64 Sequence 64, Appl
43 10 76.9 5 9 US-09-976-736-68 Sequence 68, Appl
44 10 76.9 5 9 US-09-264-516A-53 Sequence 53, Appl
45 10 76.9 5 9 US-09-264-516A-54 Sequence 54, Appl

ALIGNMENTS

RESULT 1
US-09-742-096-17
; Sequence 17, Application US/09742096
; Patent No. US20020155441A1
; GENERAL INFORMATION:

; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773050DIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide

US-09-742-096-17

Query Match 76.9%; Score 10; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
DB 2 EE 3

RESULT 2
US-09-742-096-18
; Sequence 18, Application US/09742096
; Patent No. US20020155441A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE

```

; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USODIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-09-742-096-18

```

```

Query Match          76.9%; Score 10; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches      2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 EE 5
        ||
DB      2 EE 3

```

```

RESULT 3
US-09-742-096-19
; Sequence 19, Application US/09742096
; Patent No. US20020155441A1
; GENERAL INFORMATION:
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 200773USODIV
; CURRENT APPLICATION NUMBER: US/09/742,096
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/973,642
; PRIOR FILING DATE: 1998-02-06
; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-09-742-096-19

```

```

Query Match          76.9%; Score 10; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches      2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 EE 5
        ||
DB      2 EE 3

```

```

RESULT 4
US-09-178-286-21
; Sequence 21, Application US/09178286
; Patent No. US20020168338A1
; GENERAL INFORMATION:
; APPLICANT: Baird, Andrew
; APPLICANT: Gonzalez, Ana Maria

```

```

; APPLICANT: Berry, Martin
; APPLICANT: Logan, Ann
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DELIVERY OF
; TITLE OF INVENTION: AGENTS FOR NEURONAL REGENERATION AND SURVIVAL
; FILE REFERENCE: 760100.433C1
; CURRENT APPLICATION NUMBER: US/09/178,286
; CURRENT FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide - example
; OTHER INFORMATION: cytoplasm-translocation signal sequence
US-09-178-286-21

```

```

Query Match          76.9%; Score 10; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches      2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 EE 5
        ||
DB      2 EE 3

```

```

RESULT 5
US-09-264-516A-51
; Sequence 51, Application US/09264516A
; Patent No. US20020169106A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Byers, Stephen
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR INHIBITING
; TITLE OF INVENTION: CANCER METASTASIS
; FILE REFERENCE: 100086.407C3
; CURRENT APPLICATION NUMBER: US/09/264,516A
; CURRENT FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/234,395
; PRIOR FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: 09/187,859
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/073,040
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized peptide with an OB-cadherin cell
; OTHER INFORMATION: adhesion recognition sequence
US-09-264-516A-51

```

```

Query Match          76.9%; Score 10; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches      2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      4 EE 5
        ||
DB      2 EE 3

```

```

RESULT 6
US-09-264-516A-52
; Sequence 52, Application US/09264516A
; Patent No. US20020169106A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Byers, Stephen

```

```

; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR INHIBITING
; FILE OF INVENTION: CANCER METASTASIS
; FILE REFERENCE: 100086.407C3
; CURRENT APPLICATION NUMBER: US/09/264,516A
; CURRENT FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/234,395
; PRIOR FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: 09/187,859
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/073,040
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized peptide with an OB-cadherin cell
US-09-264-516A-52

```

```

Query Match          76.9%; Score 10; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 4 EE 5
   ||
Db 1 EE 2

```

```

RESULT 7
US-09-880-132-56
; Sequence 56, Application US/09880132
; Patent No. US20020173049A1
; GENERAL INFORMATION:
; APPLICANT: Kenten, John
; APPLICANT: Roberts, Steven
; TITLE OF INVENTION: CONTROLLING PROTEIN LEVELS IN EUKARYOTIC ORGANISMS
; FILE REFERENCE: 2757-6
; CURRENT APPLICATION NUMBER: US/09/880,132
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 09/406,781
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: 60/119,851
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: binding peptide
US-09-880-132-56

```

```

Query Match          76.9%; Score 10; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 4 EE 5
   ||
Db 2 EE 3

```

```

RESULT 8
US-10-061-395-37
; Sequence 37, Application US/10061395
; Publication No. US20020192675A1
; GENERAL INFORMATION:
; APPLICANT: Zauderer, Maurice
; APPLICANT: Smith, Ernest S.

```

```

; TITLE OF INVENTION: Methods of Identifying Regulator Molecules
; FILE REFERENCE: 1821.0080003
; CURRENT APPLICATION NUMBER: US/10/061,395
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/271,423
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/265,880
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/265,589
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Heterologous signal sequence for the endoplasmic reticulum
US-10-061-395-37

```

```

Query Match          76.9%; Score 10; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 4 EE 5
   ||
Db 2 EE 3

```

```

RESULT 9
US-10-117-641-37
; Sequence 37, Application US/10117641
; Publication No. US20020194640A1
; GENERAL INFORMATION:
; APPLICANT: Misra, Santosh et al.
; TITLE OF INVENTION: PLANT PROMOTER DERIVED FROM LUMINAL BINDING PROTEIN GENE AN
; FILE OF INVENTION: ITS USE
; FILE REFERENCE: 62586
; CURRENT APPLICATION NUMBER: US/10/117,641
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 09/632,538
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Pseudotsuga menziesii
US-10-117-641-37

```

```

Query Match          76.9%; Score 10; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 4 EE 5
   ||
Db 2 EE 3

```

```

RESULT 10
US-10-087-905-5
; Sequence 5, Application US/10087905
; Publication No. US20030022152A1
; GENERAL INFORMATION:
; APPLICANT: Haaland, Perry D.
; APPLICANT: Sherman, Douglas B.
; APPLICANT: Stewart II, Walter W.
; APPLICANT: Lloyd, Sheila A.
; APPLICANT: Campbell, Robert L.
; TITLE OF INVENTION: METHODS, APPARATUS AND COMPUTER PROGRAM PRODUCTS FOR
; FILE REFERENCE: P3250
; CURRENT APPLICATION NUMBER: US/10/087,905

```

; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US/09/359,260
; PRIOR FILING DATE: 1999-07-22
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: hypothetical
; OTHER INFORMATION: peptide
US-10-087-905-5

Query Match 76.9%; Score 10; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 1 EE 2

RESULT 11
US-10-059-720-62
; Sequence 62, Application US/10059720
; Publication No. US20030027314A1
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; KRYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS WORD 97
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/059,720
; FILING DATE: 29-Jan-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: 08/890,011
; FILING DATE: 31-JULY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 62:
US-10-059-720-62

Query Match 76.9%; Score 10; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 3 EE 4

RESULT 12
US-10-235-552-3
; Sequence 3, Application US/10235552
; Publication No. US20030027768A1
; GENERAL INFORMATION:
; APPLICANT: Mazat, Andrew P.
; APPLICANT: Jones, Terence L.
; TITLE OF INVENTION: ANTI-INVASIVE AND ANTI-ANGIOGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: 38369-183655
; CURRENT APPLICATION NUMBER: US/10/235,552
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 08/900,327
; PRIOR FILING DATE: 1997-07-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: substitution,
; OTHER INFORMATION: deletion or addition variants of SEQ ID NO:2
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)
; OTHER INFORMATION: amino acid to which is bound an amino terminal
; OTHER INFORMATION: group: X attached to Pro is a peptidomimetic
; OTHER INFORMATION: compound
US-10-235-552-3

Query Match 76.9%; Score 10; DB 9; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 3 EE 4

RESULT 13
US-09-736-611-15
; Sequence 15, Application US/09736611
; Patent No. US20010023069A1
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas
; APPLICANT: Ludvigsen, Svend
; APPLICANT: Kaarsholm, Niels
; TITLE OF INVENTION: Method For Making Insulin Precursors and
; Insulin Precursor Analogs
; FILE REFERENCE: 6058.200-US
; CURRENT APPLICATION NUMBER: US/09/736,611
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 60/181,443
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/211,441
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PA 1999 01868
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: PA 2000 00440
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 4

; TYPE: PRT
; ORGANISM: N-terminal extension
US-09-736-611-15

Query Match 76.9%; Score 10; DB 10; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 1 EE 2

RESULT 14

US-09-740-359-1
; Sequence 1, Application US/09740359
; Patent No. US20010041787A1
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas Borglum
; TITLE OF INVENTION: Method for making insulin precursors and
; TITLE OF INVENTION: insulin precursor analogues having improved fermentation
; TITLE OF INVENTION: yield in yeast
; FILE REFERENCE: 6148.200-US
; CURRENT APPLICATION NUMBER: US/09/740,359
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: PA 2000 00443
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: PA 1999 01869
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/211,081
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: 60/181,450
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Mini C-peptide
US-09-740-359-1

Query Match 76.9%; Score 10; DB 10; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 1 EE 2

RESULT 15

US-09-866-824A-1
; Sequence 1, Application US/09866824A
; Patent No. US20020035243A1
; GENERAL INFORMATION:
; APPLICANT: Imfeld, Dominik
; APPLICANT: Ludin, Christian
; APPLICANT: Schreier, Thomas
; TITLE OF INVENTION: Transport System Conjugates
; FILE REFERENCE: 3006-039
; CURRENT APPLICATION NUMBER: US/09/866,824A
; CURRENT FILING DATE: 2001-09-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Transport System Conjugate
US-09-866-824A-1

Query Match 76.9%; Score 10; DB 10; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 1 EE 2

Search completed: February 11, 2003, 18:36:17
Job time : 3.42268 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:17:01 ; Search time 34.3299 Seconds
(without alignments)
93.903 Million cell updates/sec

Title: US-09-380-738a-5
Perfect score: 13
Sequence: 1 XXXEE 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	10	76.9	3	1	PCT-US01-01786A-303
2	10	76.9	3	4	US-08-017-493-21
3	10	76.9	3	20	US-09-692-077B-11
4	10	76.9	3	23	US-09-920-306-59
5	10	76.9	3	23	US-09-920-306-56
6	10	76.9	3	24	US-10-001-073-11

7	10	76.9	4	1	PCT-US00-31721-1	Sequence 1, Appl
8	10	76.9	4	1	PCT-US01-17943-94	Sequence 94, Appl
9	10	76.9	4	1	PCT-US02-02814-37	Sequence 37, Appl
10	10	76.9	4	1	PCT-US02-14457-78	Sequence 78, Appl
11	10	76.9	4	1	PCT-US93-10999-27	Sequence 27, Appl
12	10	76.9	4	1	PCT-US93-10999-29	Sequence 29, Appl
13	10	76.9	4	1	PCT-US97-04635-22	Sequence 22, Appl
14	10	76.9	4	1	PCT-US97-12618-39	Sequence 39, Appl
15	10	76.9	4	1	PCT-US97-12652-39	Sequence 39, Appl
16	10	76.9	4	1	PCT-US98-02766-93	Sequence 93, Appl
17	10	76.9	4	1	PCT-US98-04368-59	Sequence 59, Appl
18	10	76.9	4	1	PCT-US98-21231-16	Sequence 16, Appl
19	10	76.9	4	1	PCT-US98-21231-40	Sequence 40, Appl
20	10	76.9	4	1	PCT-US99-05250-70	Sequence 70, Appl
21	10	76.9	4	1	PCT-US99-16423-16	Sequence 16, Appl
22	10	76.9	4	1	PCT-US99-16423-40	Sequence 40, Appl
23	10	76.9	4	1	PCT-US99-16423-73	Sequence 73, Appl
24	10	76.9	4	3	US-07-833-836A-26	Sequence 26, Appl
25	10	76.9	4	3	US-07-912-965-134	Sequence 134, Appl
26	10	76.9	4	3	US-07-916-939-19	Sequence 19, Appl
27	10	76.9	4	3	US-07-959-949A-3	Sequence 3, Appl
28	10	76.9	4	3	US-07-959-949A-7	Sequence 7, Appl
29	10	76.9	4	4	US-08-033-883-49	Sequence 49, Appl
30	10	76.9	4	4	US-08-056-032-3	Sequence 3, Appl
31	10	76.9	4	5	US-08-134-558-3	Sequence 3, Appl
32	10	76.9	4	5	US-08-134-558-7	Sequence 7, Appl
33	10	76.9	4	6	US-08-224-831A-20	Sequence 20, Appl
34	10	76.9	4	6	US-08-224-831A-20	Sequence 20, Appl
35	10	76.9	4	6	US-08-286-059-14	Sequence 14, Appl
36	10	76.9	4	7	US-08-300-928A-84	Sequence 84, Appl
37	10	76.9	4	7	US-08-300-928A-84	Sequence 84, Appl
38	10	76.9	4	7	US-08-323-170A-3	Sequence 3, Appl
39	10	76.9	4	7	US-08-369-643-43	Sequence 43, Appl
40	10	76.9	4	7	US-08-369-643-43	Sequence 43, Appl
41	10	76.9	4	7	US-08-385-207-18	Sequence 18, Appl
42	10	76.9	4	7	US-08-385-207-18	Sequence 18, Appl
43	10	76.9	4	7	US-08-398-139B-41	Sequence 41, Appl
44	10	76.9	4	8	US-08-430-944-84	Sequence 84, Appl
45	10	76.9	4	8	US-08-438-190-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
PCT-US01-01786A-303
; Sequence 303, Application PC/TUS0101786A
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; APPLICANT: SMITHKLINE BEECHAM P.L.C.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR PERFORMING
; TITLE OF INVENTION: ANTIMICROBIAL COMPOUND SCREENING
; FILE REFERENCE: GMS0067
; CURRENT APPLICATION NUMBER: PCT/US01/01786A
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/176,870
; PRIOR FILING DATE: 2000-01-18
; NUMBER OF SEQ IDS NOS: 403
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 303
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Escherichia coli
PCT-US01-01786A-303

Query Match	Best Local	Similarity	Score	DB 1;	Length	3;
Matches	2;	Conservative	100.0%;	Pred. No. 4.2e+06;	0;	Indels
				Mismatches	0;	Gaps
					0;	Gaps
Qy	4	EE	5			
Db	2	EE	3			

```

Db      1  EE 2

RESULT 4
US-09-920-306-59
; Sequence 59, Application US/09920306
; GENERAL INFORMATION:
; APPLICANT: Unilever PLC
; APPLICANT: Unilever NV
; TITLE OF INVENTION: Peptides Capable of Functioning as Mimotopes for
; FILE REFERENCE: Peptide Mimotopes
; TITLE OF INVENTION: Hormonal Analytes
; CURRENT APPLICATION NUMBER: US/09/920,306
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: EP00306613.1
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-920-306-59

Query Match      76.9%; Score 10; DB 23; Length 3;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4  EE 5
        1  1
Db      2  EE 3

RESULT 5
US-09-920-306-66
; Sequence 66, Application US/09920306
; GENERAL INFORMATION:
; APPLICANT: Unilever PLC
; APPLICANT: Unilever NV
; TITLE OF INVENTION: Peptides Capable of Functioning as Mimotopes for
; FILE REFERENCE: Peptide Mimotopes
; TITLE OF INVENTION: Hormonal Analytes
; CURRENT APPLICATION NUMBER: US/09/920,306
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: EP00306613.1
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-920-306-66

Query Match      76.9%; Score 10; DB 23; Length 3;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4  EE 5
        1  1
Db      1  EE 2

RESULT 6
US-10-001-073-11
; Sequence 11, Application US/10001073

```

```

RESULT 2
US-08-017-493-21
; Sequence 21, Application US/08017493
; GENERAL INFORMATION:
; APPLICANT: Chandrasegaran, Srinivasan
; TITLE OF INVENTION: Functional Domains in Foki Restriction
; TITLE OF INVENTION: Endonuclease
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman, Darby & Cushman
; STREET: 1100 New York Ave., N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/017,493
; FILING DATE: 19930212
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kokulis, Paul N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: PNK/4130/97635/CLB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3503
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-017-493-21

Query Match      76.9%; Score 10; DB 4; Length 3;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4  EE 5
        1  1
Db      1  EE 2

RESULT 3
US-09-692-077B-11
; Sequence 11, Application US/09692077B
; GENERAL INFORMATION:
; APPLICANT: Liggett, Stephen
; TITLE OF INVENTION: Alpha-2B-adrenergic receptor polymorphisms
; FILE REFERENCE: Sequences 1-22
; CURRENT APPLICATION NUMBER: US/09/692,077B
; CURRENT FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-692-077B-11

Query Match      76.9%; Score 10; DB 20; Length 3;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4  EE 5
        1  1

```



```

; GENERAL INFORMATION:
; APPLICANT: Liggett, Stephen
; TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
; FILE REFERENCE: 13073-PCT
; CURRENT APPLICATION NUMBER: US/10/001.073
; CURRENT FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 3
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-001-073-11

Query Match          76.9%; Score 10; DB 24; Length 3;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
Db 1 EE 2

RESULT 7
PCT-US00-31721-1
; Sequence 1, Application PC/TUS0031721
; GENERAL INFORMATION:
; APPLICANT: MURPHY, JOHN R.
; APPLICANT: HARRISON, ROBERT J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATMENT OF DISEASE
; FILE REFERENCE: AMSC 3.4-003
; CURRENT APPLICATION NUMBER: PCT/US00/31721
; CURRENT FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)
; OTHER INFORMATION: phospho-tyrosine
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Phospho-tyrosine-containing peptide
PCT-US00-31721-1

Query Match          76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
Db 2 EE 3

RESULT 8
PCT-US01-17943-94
; Sequence 94, Application PC/TUS0117943
; GENERAL INFORMATION:
; APPLICANT: Haaland, Perry D.
; APPLICANT: Sherman, Douglas
; APPLICANT: Campbell, Robert L.
; APPLICANT: Stewart, William
; APPLICANT: Lloyd, Sheila A.
; APPLICANT: Erickson, Bruce W. (deceased)
; TITLE OF INVENTION: Peptides for Use in Culture Media
; FILE REFERENCE: Peptides for Culture Media
; CURRENT APPLICATION NUMBER: PCT/US01/17943
; CURRENT FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 94
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide linker
PCT-US01-17943-94

```

```

Query Match          76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 EE 5
Db 2 EE 3

```

```

RESULT 9
PCT-US02-02814-37
; Sequence 37, Application PC/TUS0202814
; GENERAL INFORMATION:
; APPLICANT: University of Rochester
; APPLICANT: Zauderer, Maurice (U.S. Only)
; APPLICANT: Smith, Ernest S. (U.S. Only)
; TITLE OF INVENTION: Methods of Identifying Regulator Molecules
; FILE REFERENCE: 1821.0080003
; CURRENT APPLICATION NUMBER: PCT/US02/02814
; CURRENT FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: 60/271,423
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/265,880
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/265,589
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Heterologous signal sequence for the endoplasmic reticulum
PCT-US02-02814-37

```

```

Query Match          76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 EE 5
Db 2 EE 3

```

```

RESULT 10
PCT-US02-14457-78
; Sequence 78, Application PC/TUS0214457
; GENERAL INFORMATION:
; APPLICANT: AlphaGene, Inc.
; APPLICANT: Stark, Karen A.
; APPLICANT: Weaver, Alix
; APPLICANT: Hoffmann, Heidi M.
; APPLICANT: Krauss, Raul
; APPLICANT: Valenzuela, Dario B.
; APPLICANT: Saini, Kulvinder Singh
; TITLE OF INVENTION: Cell Adhesion-Mediating Proteins and
; FILE REFERENCE: Polynucleotides Encoding Them
; FILE REFERENCE: 1966.1014002
; CURRENT APPLICATION NUMBER: PCT/US02/14457
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 60/289,179
; PRIOR FILING DATE: 2001-05-07
; PRIOR APPLICATION NUMBER: 60/315,736

```

us-09-380-738a-5.rapm

Wed Feb 12 11:35:25 2003

```

; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SH2 domain
PCT-US02-14457-78

Query Match          76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
Db 2 EE 3

RESULT 11
PCT-US93-10999-27
; Sequence 27, Application PC/TUS9310999
; GENERAL INFORMATION:
; APPLICANT: Jameson, Bradford A.
; APPLICANT: McDonnell, James M.
; APPLICANT: Korngold, Robert
; TITLE OF INVENTION: Compounds That Inhibit T Cell Proliferation And
; TITLE OF INVENTION: Methods Using The Same
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb Storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10999
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/076,092
; FILING DATE: 11-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/977,692
; FILING DATE: 13-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-904
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4
; TYPE: Amino Acid
; TOPOLOGY: linear
PCT-US93-10999-29

Query Match          76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
Db 2 EE 3

RESULT 13
PCT-US97-04635-22
; Sequence 22, Application PC/TUS9704635
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC.
; APPLICANT: MOREHOUSE SCHOOL OF MEDICINE
; APPLICANT: POTEPA, JAN.
; APPLICANT: TRAVIS, JAMES
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS COMPRISING
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS PROTEINS AND/OR PEPTIDES AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201

```

;; CITY: Boulder
;; STATE: CO
;; COUNTRY: US
;; ZIP: 80303
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US97/04635
;; FILING DATE: 21-MAR-1997
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 60/013,945
;; FILING DATE: 22-MAR-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ferber, Donna M.
;; REGISTRATION NUMBER: 33,878
;; REFERENCE/DOCKET NUMBER: 103-95
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (303) 488-8080
;; TELEFAX: (303) 499-8089
;; INFORMATION FOR SEQ ID NO: 22:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
PCT-US97-04635-22

Query Match 76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0;

QY 4 EE 5
Db 2 EE 3

RESULT 14
PCT-US97-12618-39
;; Sequence 39, Application PC/TUS9712618
;; GENERAL INFORMATION:
;; APPLICANT: University of Utah Research Foundation
;; TITLE OF INVENTION: Cognetix, Inc.
;; NUMBER OF SEQUENCES: 71
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
;; STREET: 1201 New York Avenue, N.W., Suite 1000
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US97/12618
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/684,742
;; FILING DATE: 22-JUL-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ihnen, Jeffrey L.
;; REGISTRATION NUMBER: 28,957
;; REFERENCE/DOCKET NUMBER: 24260-121389-WO2

;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-962-4800
;; TELEFAX: 202-962-8300
;; INFORMATION FOR SEQ ID NO: 39:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FRAGMENT TYPE: C-terminal
PCT-US97-12618-39

Query Match 76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0;

QY 4 EE 5
Db 2 EE 3

RESULT 15
PCT-US97-12652-39
;; Sequence 39, Application PC/TUS9712652
;; GENERAL INFORMATION:
;; APPLICANT: Cognetix, Inc.
;; TITLE OF INVENTION: Use of Conantokins
;; NUMBER OF SEQUENCES: 71
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
;; STREET: 1201 New York Avenue, N.W., Suite 1000
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US97/12652
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/762,377
;; FILING DATE: 06-DEC-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/684,750
;; FILING DATE: 22-JUL-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ihnen, Jeffrey L.
;; REGISTRATION NUMBER: 28,957
;; REFERENCE/DOCKET NUMBER: 24260-121389-WO1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-962-4800
;; TELEFAX: 202-962-8300
;; INFORMATION FOR SEQ ID NO: 39:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 4 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FRAGMENT TYPE: C-terminal
PCT-US97-12652-39

Query Match 76.9%; Score 10; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 2; Conservative 0; Mismatches 0; Indels 0;

QY 4 EE 5

Wed Feb 12 11:35:25 2003

us-09-380-738a-5.rapm

Page 6

Db 11
 2 EE 3

Search completed: February 11, 2003, 18:33:26
Job time : 35.3299 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:17:56 ; Search time 5.72165 Seconds
(without alignments)
72.734 Million cell updates/sec

Title: US-09-380-738A-5
Perfect score: 13
Sequence: 1 XXXEE 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 420401 seqs, 83231269 residues

Total number of hits satisfying chosen parameters: 420401

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	10	76.9	3	6	US-10-275-360-15
2	10	76.9	4	1	PCT-US02-32007-57
3	10	76.9	4	1	PCT-US02-00667A-72
4	10	76.9	4	5	US-09-783-130A-8
5	10	76.9	4	5	US-09-611-257A-45
6	10	76.9	4	5	US-09-611-257A-47
7	10	76.9	4	5	US-09-856-886B-16
8	10	76.9	4	5	US-09-165-062A-28
9	10	76.9	4	5	US-09-799-250A-715
10	10	76.9	4	6	US-10-062-109A-721
11	10	76.9	4	6	US-10-316-421-15
12	10	76.9	4	6	US-10-191-254C-3
13	10	76.9	4	6	US-10-099-408A-10
14	10	76.9	4	6	US-10-176-791A-21
15	10	76.9	4	6	US-10-345-281-56
16	10	76.9	4	6	US-10-319-003-99
17	10	76.9	4	6	US-10-005-480A-721
18	10	76.9	5	1	PCT-US01-32150-90
19	10	76.9	5	5	US-09-653-812B-138
20	10	76.9	5	5	US-09-992-124B-13
21	10	76.9	5	5	US-09-992-124B-57
22	10	76.9	5	5	US-09-992-124B-62
23	10	76.9	5	5	US-09-800-770-50
24	10	76.9	5	5	US-09-800-770-53
25	10	76.9	5	5	US-09-636-243B-65
26	10	76.9	5	5	US-09-856-886B-17

Sequence 25, Appl
Sequence 26, Appl
Sequence 15, Appl
Sequence 11, Appl
Sequence 32, Appl
Sequence 1, Appl
Sequence 100, Appl
Sequence 21, Appl
Sequence 91, Appl
Sequence 10, Appl
Sequence 5, Appl
Sequence 79, Appl
Sequence 1, Appl
Sequence 5, Appl
Sequence 70, Appl
Sequence 70, Appl
Sequence 18, Appl

5 5 US-09-856-886B-25
5 5 US-09-856-886B-26
5 6 US-10-281-652-15
5 6 US-10-303-689-11
5 6 US-10-213-512-32
5 6 US-10-336-839-1
5 6 US-10-319-003-100
5 6 US-10-355-208-20
5 6 US-10-355-208-21
6 1 PCT-US01-32150-91
6 4 US-08-477-504D-10
6 5 US-09-856-886-5
6 5 US-09-668-314C-79
6 5 US-09-749-959-1
6 5 US-09-776-268A-5
6 5 US-09-776-268A-6
6 5 US-09-794-925A-70
6 5 US-09-794-927A-70
6 5 US-09-368-632B-18

ALIGNMENTS

RESULT 1

US-10-275-360-15

; Sequence 15, Application US/10275360

; GENERAL INFORMATION:

; APPLICANT: JOMAA, Hassan

; TITLE OF INVENTION: GENES OF THE 1-DESOXY-D-XYLULOSE BIOSYNTHESIS PATH

; FILE REFERENCE: JOMAA-12 (PCT)

; CURRENT APPLICATION NUMBER: US/10/275,360

; PRIOR FILING DATE: 2002-11-05

; PRIOR APPLICATION NUMBER: PCT/EP01/04537

; PRIOR FILING DATE: 2001-04-21

; PRIOR APPLICATION NUMBER: DE10021688.9

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 15

; LENGTH: 3

; TYPE: PRT

; ORGANISM: Plasmodium falciparum

US-10-275-360-15

Query Match 76.9%; Score 10; DB 6; Length 3;

Best Local Similarity 100.0%; Pred. No. 3.5e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 EE 5

Db 2 EE 3

RESULT 2

PCT-US02-32007-57

; Sequence 57, Application PC/TUS0232007

; GENERAL INFORMATION:

; APPLICANT: Yale University

; TITLE OF INVENTION: NOGO RECEPTOR-MEDIATED BLOCKADE OF AXONAL GROWTH

; FILE REFERENCE: C077 CIP PCT

; CURRENT APPLICATION NUMBER: PCT/US02/32007

; CURRENT FILING DATE: 2002-10-04

; PRIOR APPLICATION NUMBER: 09/972,599

; PRIOR FILING DATE: 2001-10-06

; NUMBER OF SEQ ID NOS: 57

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 57

; LENGTH: 4

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

; OTHER INFORMATION: peptide

PCT-US02-32007-57

Query Match 76.9%; Score 10; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.5e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 1 EE 2

RESULT 3

PCT-US02-00667A-72

; Sequence 72, Application PC/TUS0200667A
 ; GENERAL INFORMATION:
 ; APPLICANT: UNIVERSITY OF MARYLAND, COLLEGE PARK
 ; TITLE OF INVENTION: METHODS FOR DETERMINING RING NUMBER IN CAROTENOIDS BY
 ; FILE REFERENCE: LYCOPENE EPSILON - CYCLES AND USES THEREOF
 ; CURRENT APPLICATION NUMBER: 108172-00055
 ; CURRENT FILING DATE: 2002-09-30
 ; PRIOR FILING DATE: 2001-01-12
 ; NUMBER OF SEQ ID NOS: 97
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 72
 ; LENGTH: 4
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Illustrative
 ; OTHER INFORMATION: peptide junction sequence

PCT-US02-00667A-72

Query Match 76.9%; Score 10; DB 1; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.5e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 1 EE 2

RESULT 4

US-09-783-130A-8

; Sequence 8, Application US/09783130A
 ; GENERAL INFORMATION:
 ; APPLICANT: WEBSTER, Marion
 ; APPLICANT: et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
 ; FILE REFERENCE: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
 ; CURRENT APPLICATION NUMBER: US/09/783,130A
 ; CURRENT FILING DATE: 2001-02-15
 ; PRIOR FILING DATE: 2000-12-19
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 4
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-09-783-130A-8

Query Match 76.9%; Score 10; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.5e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 3 EE 4

RESULT 5

US-09-611-257A-45

; Sequence 45, Application US/09611257A
 ; GENERAL INFORMATION:
 ; APPLICANT: Snutch, Terrance
 ; APPLICANT: Baillie, David L.
 ; TITLE OF INVENTION: MAMMALIAN T-TYPE CALCIUM CHANNELS
 ; FILE REFERENCE: 38109-20007.21
 ; CURRENT APPLICATION NUMBER: US/09/611,257A
 ; CURRENT FILING DATE: 2000-07-06
 ; PRIOR FILING DATE: 1999-07-02
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR FILING DATE: 1997-02-28
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 45
 ; LENGTH: 4
 ; TYPE: PRT
 ; ORGANISM: Unknown
 ; FEATURE:
 ; OTHER INFORMATION: T-type channels in P-regions of domains I-IV
 ; OTHER INFORMATION: T-type channels in P-regions of domains I-IV

US-09-611-257A-45

Query Match 76.9%; Score 10; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.5e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 1 EE 2

RESULT 6

US-09-611-257A-47

; Sequence 47, Application US/09611257A
 ; GENERAL INFORMATION:
 ; APPLICANT: Snutch, Terrance
 ; APPLICANT: Baillie, David L.
 ; TITLE OF INVENTION: MAMMALIAN T-TYPE CALCIUM CHANNELS
 ; FILE REFERENCE: 38109-20007.21
 ; CURRENT APPLICATION NUMBER: US/09/611,257A
 ; CURRENT FILING DATE: 2000-07-06
 ; PRIOR FILING DATE: 1999-07-02
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR FILING DATE: 1997-02-28
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 47
 ; LENGTH: 4
 ; TYPE: PRT
 ; ORGANISM: Unknown
 ; FEATURE:
 ; OTHER INFORMATION: L-type calcium channels in P-regions of domains
 ; OTHER INFORMATION: I-IV

US-09-611-257A-47

Query Match 76.9%; Score 10; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.5e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 1 EE 2

RESULT 7
 US-09-856-886B-16
 ; Sequence 16, Application US/09856886B
 ; GENERAL INFORMATION:
 ; APPLICANT: Bianchi, Elisabetta
 ; APPLICANT: Ingallinella, Paolo
 ; APPLICANT: Pessi, Antonello
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOUNDS FOR THE
 ; FILE REFERENCE: IT0013P
 ; CURRENT APPLICATION NUMBER: US/09/856,886B
 ; CURRENT FILING DATE: 2001-07-26
 ; PRIOR APPLICATION NUMBER: PCT/EP99/09207
 ; PRIOR FILING DATE: 1999-11-24
 ; PRIOR APPLICATION NUMBER: GB/9825946.8
 ; PRIOR FILING DATE: 1998-11-26
 ; NUMBER OF SEQ ID NOS: 109
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 16
 ; LENGTH: 4
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic sequence
 ; FEATURE:
 ; NAME/KEY: MOD_RES
 ; LOCATION: 1
 ; OTHER INFORMATION: Acetylation
 US-09-856-886B-16

Query Match 76.9%; Score 10; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.5e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 Db 2 EE 3

RESULT 8
 US-09-165-062A-28
 ; Sequence 28, Application US/09165062A
 ; GENERAL INFORMATION:
 ; APPLICANT: LEV, SIMA
 ; TITLE OF INVENTION: SCHLESINGER, JOSEPH
 ; FILE REFERENCE: 038602/0116
 ; CURRENT APPLICATION NUMBER: US/09/165,062A
 ; CURRENT FILING DATE: 1998-10-01
 ; PRIOR APPLICATION NUMBER: 08/460,626
 ; PRIOR FILING DATE: 1995-06-02
 ; PRIOR APPLICATION NUMBER: 08/357,642
 ; PRIOR FILING DATE: 1994-12-15
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 28
 ; LENGTH: 4
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-165-062A-28

Query Match 76.9%; Score 10; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.5e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 Db 2 EE 3

RESULT 9
 US-09-799-250A-715
 ; Sequence 715, Application US/09799250A

; GENERAL INFORMATION:
 ; APPLICANT: Pia M. Challita-Eid
 ; APPLICANT: Rene S. Hubert
 ; APPLICANT: Steve Chappell Mitchell
 ; APPLICANT: Arthur B. Raitano
 ; APPLICANT: Mary Faris
 ; APPLICANT: Daniel E.H. Afar
 ; APPLICANT: Aya Jakobovits
 ; TITLE OF INVENTION: 12IP1F1: A TISSUE SPECIFIC PROTEIN
 ; FILE REFERENCE: 129.34US01 (51158-20034.00)
 ; CURRENT APPLICATION NUMBER: US/09/799,250A
 ; CURRENT FILING DATE: 2001-03-05
 ; NUMBER OF SEQ ID NOS: 720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 715
 ; LENGTH: 4
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-799-250A-715

Query Match 76.9%; Score 10; DB 5; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.5e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 Db 3 EE 4

RESULT 10
 US-10-062-109A-721
 ; Sequence 721, Application US/10062109A
 ; GENERAL INFORMATION:
 ; APPLICANT: Agensys
 ; APPLICANT: Challita-Eid, Pia M.
 ; APPLICANT: Raitano, Arthur B.
 ; APPLICANT: Faris, Mary
 ; APPLICANT: Hubert, Rene S.
 ; APPLICANT: Morrison, Karen Jane Meyrick
 ; APPLICANT: Jakobovits, Aya
 ; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
 ; TITLE OF INVENTION: Entitled 161P2F10B Useful in Treatment and Detection of
 ; FILE REFERENCE: 51158-20062.01
 ; CURRENT APPLICATION NUMBER: US/10/062,109A
 ; CURRENT FILING DATE: 2002-01-31
 ; PRIOR APPLICATION NUMBER: US 10/005,480
 ; PRIOR FILING DATE: 2001-11-07
 ; NUMBER OF SEQ ID NOS: 765
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 721
 ; LENGTH: 4
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-062-109A-721

Query Match 76.9%; Score 10; DB 6; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.5e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 Db 3 EE 4

RESULT 11
 US-10-316-421-15
 ; Sequence 15, Application US/10316421
 ; GENERAL INFORMATION:
 ; APPLICANT: Kjeldsen, Thomas
 ; APPLICANT: Ludvigsen, Svend
 ; APPLICANT: Kaarsholm, Niels

; TITLE OF INVENTION: Method For Making Insulin Precursors and
; TITLE OF INVENTION: Insulin Precursor Analogs
; FILE REFERENCE: 6058,200-US
; CURRENT APPLICATION NUMBER: US/10/316,421
; CURRENT FILING DATE: 2002-12-11
; PRIOR FILING DATE: 2000-12-14
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 60/181,443
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/211,441
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PA 1999 01868
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: PA 2000 00440
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 4
; TYPE: PRT
; ORGANISM: N-terminal extension
US-10-316-421-15

Query Match 76.9%; Score 10; DB 6; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 1 EE 2

RESULT 12

US-10-191-254C-3
; Sequence 3, Application US/10191254C

; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: CASPASE 9 ACTIVATION AND USES THEREFOR
; FILE REFERENCE: AM101006
; CURRENT APPLICATION NUMBER: US/10/191,254C
; CURRENT FILING DATE: 2002-12-03
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Substrate
US-10-191-254C-3

Query Match 76.9%; Score 10; DB 6; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 2 EE 3

RESULT 13

US-10-099-408A-10
; Sequence 10, Application US/10099408A

; GENERAL INFORMATION:
; APPLICANT: D'Lima, Darryl
; APPLICANT: Lotz, Martin
; APPLICANT: Colwell, Clifford
; TITLE OF INVENTION: Process of Inhibiting Cell Death in
; FILE REFERENCE: TSRI 801.1/NOV 0237P
; CURRENT APPLICATION NUMBER: US/10/099,408A
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/276,183

; PRIOR FILING DATE: 2001-03-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: caspase inhibitor; fluoromethyl ketone on the
; OTHER INFORMATION: C-terminus
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: 1
; OTHER INFORMATION: Carboxyloxy protection of the N-terminal amino
; OTHER INFORMATION: group
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: 4
; OTHER INFORMATION: Methyl ester protection of the carboxylic acid
; OTHER INFORMATION: side chain;
; OTHER INFORMATION: C-terminal fluoromethyl ketone where the C-1
; OTHER INFORMATION: carbon of the aspartyl residue is the carbonyl
; OTHER INFORMATION: carbon of the ketone
US-10-099-408A-10

Query Match 76.9%; Score 10; DB 6; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 2 EE 3

RESULT 14

US-10-176-791A-21
; Sequence 21, Application US/10176791A

; GENERAL INFORMATION:
; APPLICANT: WEHRLE-HALLER, BERNHARD M.
; APPLICANT: IMHOF, BEAT A.
; TITLE OF INVENTION: Basolateral Sorting Signal and
; TITLE OF INVENTION: Inhibitors Thereof
; FILE REFERENCE: 50275/002001
; CURRENT APPLICATION NUMBER: US/10/176,791A
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: PCT/EP00/13141
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: PCT/CH99/00624
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Cluster of Charged Amino-Acids
US-10-176-791A-21

Query Match 76.9%; Score 10; DB 6; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 1 EE 2

RESULT 15

US-10-345-281-56
; Sequence 56, Application US/10345281

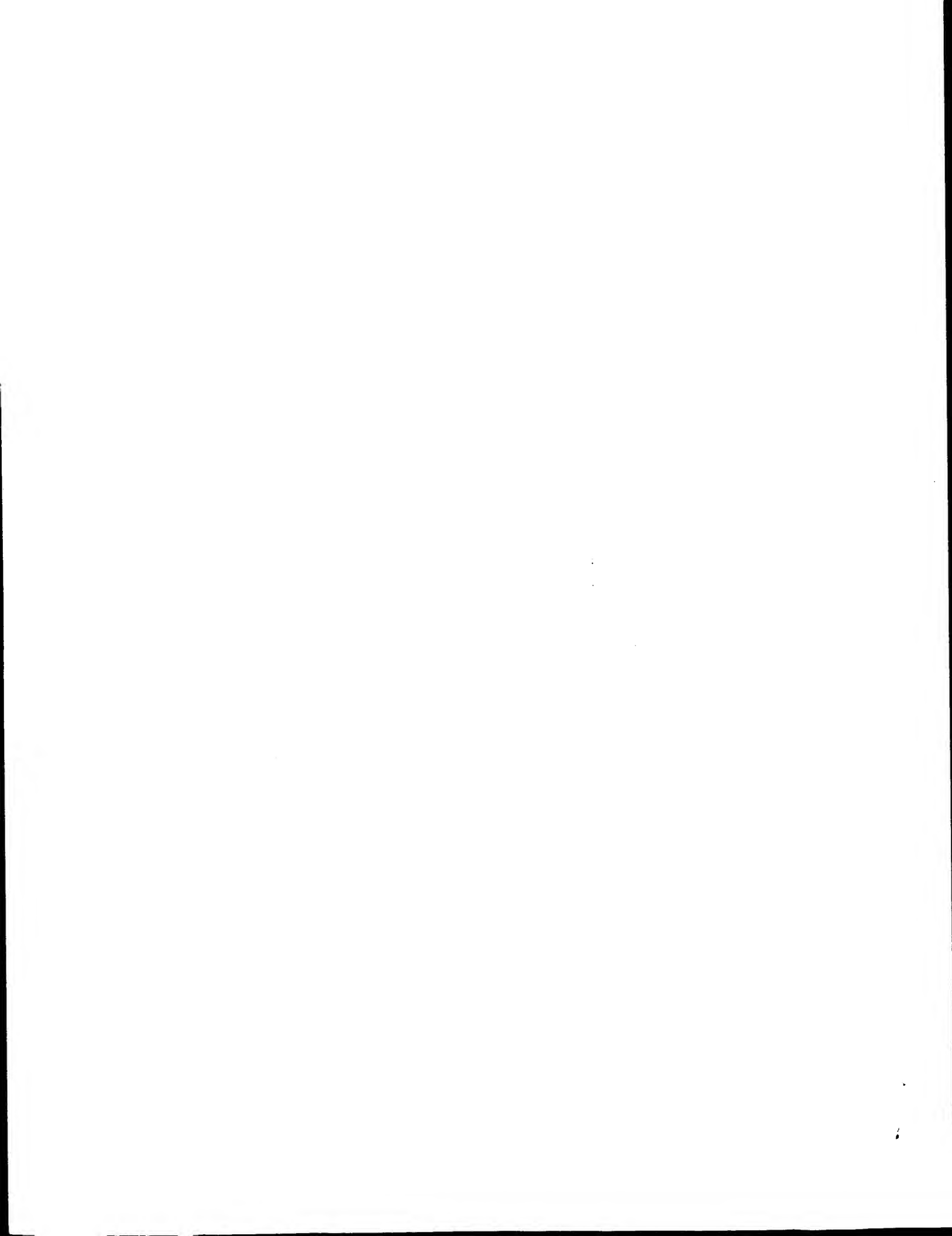
; GENERAL INFORMATION:
; APPLICANT: Kenten, John

; APPLICANT: Roberts, Steven
; TITLE OF INVENTION: CONTROLLING PROTEIN LEVELS IN EUKARYOTIC ORGANISMS
; FILE REFERENCE: 2757-6
; CURRENT APPLICATION NUMBER: US/10/345,281
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US/09/880,132
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 09/406,781
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: 60/119,851
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: binding peptide
US-10-345-281-56

Query Match 76.9%; Score 10; DB 6; Length 4;
Best Local Similarity 100.0%; Pred. No. 3.5e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 2 EE 3

Search completed: February 11, 2003, 18:35:22
Job time : 6.72165 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:14:11 ; Search time 4.38144 Seconds
(without alignments)
109.706 Million cell updates/sec

Title: US-09-380-738A-5

Perfect score: 13

Sequence: 1 XXXEE 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	DB	ID	Description
1	10	76.9	4	2	S55238		pallidipin - assas
2	10	76.9	5	2	A32014		trAM protein - Esc
3	10	76.9	6	2	S11556		hydrogensulfite re
4	10	76.9	7	2	I46868		alpha-myosin heavy
5	10	76.9	7	2	B39040		calsequestrin, fas
6	10	76.9	7	2	S45648		Na+-transporting A
7	10	76.9	8	2	S29272		tocopherol-binding
8	10	76.9	8	2	B47594		aspartate kinase (
9	10	76.9	8	2	S69165		ferredoxin a2 - Ja
10	10	76.9	9	2	A44873		caldesmon - rabbit
11	10	76.9	9	2	A61230		caldesmon - car
12	10	76.9	9	2	C24180		fibrinogen beta ch
13	10	76.9	9	2	C60070		fibrinogen beta ch
14	10	76.9	9	2	PH0108		gastrin - domestic
15	10	76.9	9	2	PH0942		late G1-69 protein
16	10	76.9	9	2	S65388		T-cell receptor be
17	10	76.9	10	2	S53388		cytochrome-c oxida
18	10	76.9	10	2	S39392		calpain (EC 3.4.22
19	10	76.9	10	2	C38925		seed storage prote
20	10	76.9	10	2	S43625		cytochrome-c oxida
21	10	76.9	11	2	PQ0662		photosystem I 17.5
22	10	76.9	11	2	B39853		LuxC protein - Pho
23	10	76.9	11	2	A55149		tetracenomycin A2
24	10	76.9	11	2	JQ2307		hypothetical 1.5K
25	10	76.9	11	2	JQ2317		hypothetical 1.5K
26	10	76.9	11	2	S65395		chemical-sense-rel
27	10	76.9	11	2	S43626		cytochrome-c oxida
28	10	76.9	11	2	A14454		6-phosphofructokin
29	10	76.9	11	2	D45900		complement C3b rec

30	10	76.9	11	2	S09349	microtubule-associ
31	10	76.9	11	4	S41909	hypothetical prote
32	10	76.9	11	4	PC2124	aminotransferase c
33	10	76.9	12	2	A39233	myosin heavy chain
34	10	76.9	12	2	T46794	hypothetical prote
35	10	76.9	12	2	A38925	seed storage prote
36	10	76.9	12	2	S16335	beta-conglycinin a
37	10	76.9	12	2	F61308	hemocyanin chain 5
38	10	76.9	12	2	B47171	chondroitin sulfat
39	10	76.9	12	2	PT0228	Ig heavy chain CDR
40	10	76.9	12	2	S68271	major urinary prot
41	10	76.9	12	2	S27024	Na+/K+-exchanging
42	10	76.9	12	2	S27023	Na+/K+-exchanging
43	10	76.9	12	4	PC2122	aminotransferase c
44	10	76.9	13	2	S39413	tubulin beta chain
45	10	76.9	13	2	PA0031	protein QA300045 -

ALIGNMENTS

RESULT 1

S55238

pallidipin - assasin bug (fragment)

C:Species: Triatoma pallidipennis (assassin bug)

C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 19-May-2000

C:Accession: S55238

R:Haendler, B.; Becker, A.; Noeske-Jungblut, C.; Kraetzschmar, J.; Donner, P.; Sch

Biochem. J. 307, 465-470, 1995

A:Title: Expression of active recombinant pallidipin, a novel platelet aggregation

A:Reference number: S55238; MUID:95251610; PMID:7733884

A:Accession: S55238

A:Molecule type: protein

A:Residues: 1-4 <HAE>

Query Match 76.9%; Score 10; DB 2; Length 4;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5

Db 1 EE 2

RESULT 2

A32014

trAM protein - Escherichia coli plasmid R100 (fragment)

C:Species: Escherichia coli

C:Date: 22-Jun-1989 #sequence_revision 22-Jun-1989 #text_change 16-Feb-1997

C:Accession: A32014

R:Inamoto, S.; Yoshioka, Y.; Ohtsubo, E.

J. Bacteriol. 170, 2749-2757, 1988

A:Title: Identification and characterization of the products from the trAJ and tra

A:Reference number: A32014; MUID:88227859; PMID:2836369

A:Accession: A32014

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-5 <INA>

C:Genetics:

A:Genome: plasmid

C:Keywords: DNA binding

Query Match 76.9%; Score 10; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5

Db 4 EE 5

RESULT 3

S11556

hydrogensulfite reductase (EC 1.8.99.3) chain 2 - Desulfovibrio thermophilus (fragment)
 N:Alternate names: bisulfite reductase; desulfosulfidase
 C:Species: Desulfovibrio thermophilus
 C:Date: 19-Mar-1997 #sequence_revision 12-Dec-1997 #text_change 30-Jan-1998
 C:Accession: S11556
 R:Faugue, G.; Lino, A.R.; Czechowski, M.; Kang, L.; DerVartanian, D.V.; Moura, J.J.G.; I Biochim. Biophys. Acta 1040, 112-118, 1990
 A:Title: Purification and characterization of bisulfite reductase (desulfosulfidase) from
 A:Reference number: S11024; MUID:90335276; PMID:2165817
 A:Accession: S11556
 A:Molecule type: protein
 A:Residues: 1-6 <FAU>
 C:Keywords: oxidoreductase

Query Match 76.9%; Score 10; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 4

I46868
 alpha-myosin heavy chain - rabbit (fragment)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 05-Nov-1999
 C:Accession: I46868
 R:Friedman, D.J.; Umeda, P.K.; Sinha, A.M.; Hsu, H. Proc. Natl. Acad. Sci. U.S.A. 81, 3044-3048, 1984
 A:Title: Characterization of genomic clones specifying rabbit alpha- and beta-ventricular
 A:Reference number: I46868; MUID:84221901; PMID:6328491
 A:Accession: I46868
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-7 <FRI>
 A:Cross-references: GB:K01698; NID:g165538; PIDN:AAA31415.1; PID:g165539

Query Match 76.9%; Score 10; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 0;

QY 4 EE 5
 ||
 Db 6 EE 7

RESULT 5

B39040
 calsequestrin, fast skeletal muscle - rabbit (fragment)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 23-Feb-1997
 C:Accession: B39040
 R:Cala, S.E.; Jones, L.R. J. Biol. Chem. 266, 391-398, 1991
 A:Title: Phosphorylation of cardiac and skeletal muscle calsequestrin isoforms by casein
 A:Reference number: A39040; MUID:91093153; PMID:1985907
 A:Accession: B39040
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-7 <CAL>
 C:Keywords: phosphoprotein; skeletal muscle

Query Match 76.9%; Score 10; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 0;

QY 4 EE 5
 ||
 Db 3 EE 4

RESULT 6

S45648
 Na⁺-transporting ATP synthase (EC 3.6.1.1) alpha chain - Acetobacterium woodii (fra
 N:Alternate names: ATPase alpha chain
 C:Species: Acetobacterium woodii
 C:Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 07-May-1999
 C:Accession: S45648
 R:Reidlinger, J.; Mueller, V. Eur. J. Biochem. 223, 275-283, 1994
 A:Title: Purification of ATP synthase from Acetobacterium woodii and identification
 A:Reference number: S45648; MUID:94307271; PMID:8033902
 A:Accession: S45648
 A:Molecule type: protein
 A:Residues: 1-3; 4-7 <REI>
 A:Experimental source: DSM 1030
 C:Keywords: hydrolase

Query Match 76.9%; Score 10; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 0;

QY 4 EE 5
 ||
 Db 5 EE 6

RESULT 7

S29272
 tocopherol-binding protein, 81k - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 22-Nov-1993 #sequence_revision 12-Apr-1996 #text_change 12-Apr-1996
 C:Accession: S29272
 R:Nalecz, K.A.; Nalecz, M.J.; Azzi, A. Eur. J. Biochem. 209, 37-42, 1992
 A:Title: Isolation of tocopherol-binding proteins from the cytosol of smooth muscle
 A:Reference number: S29272; MUID:93011150; PMID:1396710
 A:Accession: S29272
 A:Molecule type: protein
 A:Residues: 1-8 <NAL>
 A:Experimental source: smooth muscle A7r5 cells

Query Match 76.9%; Score 10; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 8

B47594
 aspartate kinase (EC 2.7.2.4) beta chain - Corynebacterium flavum (strain N13) (fra
 C:Species: Corynebacterium flavum
 C:Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 19-Dec-1997
 C:Accession: B47594
 R:Follettie, M.T.; Peoples, O.P.; Agoropoulou, C.; Sinskey, A.J. J. Bacteriol. 175, 4096-4103, 1993
 A:Title: Gene structure and expression of the Corynebacterium flavum N13 ask-
 A:Reference number: A47594; MUID:93308089; PMID:8100567
 A:Accession: B47594
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-8 <FOL>
 C:Keywords: phosphotransferase

Query Match 76.9%; Score 10; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 9

S69165
 ferredoxin a2 - Japanese radish (fragment)
 C:Species: Kaiware daikon (Japanese radish)
 C:Date: 10-Mar-1998 #sequence_revision 10-Mar-1998 #text_change 17-Apr-1998
 C:Accession: S69165
 R:Obata, S.; Nishimura, M.; Nagai, K.; Sakihama, N.; Shin, M.
 Arch. Biochem. Biophys. 316, 757-802, 1995
 A:Title: Four ferredoxins from Japanese radish leaves.
 A:Reference number: S69164; MUID:95168867; PMID:7864635
 A:Accession: S69165
 A:Molecule type: protein
 A:Residues: 1-8 <OBA>
 C:Keywords: 2Fe-2S; electron transfer; iron-sulfur protein

Query Match 76.9%; Score 10; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 4 EE 5

RESULT 10

A44873
 caldesmon - rabbit (fragment)
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 31-Oct-1997
 C:Accession: A44873
 R:Ikebe, M.; Hornick, T.
 Arch. Biochem. Biophys. 288, 538-542, 1991
 A:Title: Determination of the phosphorylation sites of smooth muscle caldesmon by protein kinase C.
 A:Reference number: A44873; MUID:91378496; PMID:1898046
 A:Accession: A44873
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-9 <IKE>
 A:Experimental source: skeletal myosin
 A:Note: sequence extracted from NCBI backbone (NCBIP:63199)
 C:Superfamily: caldesmon

Query Match 76.9%; Score 10; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 7 EE 8

RESULT 11

A61230
 calsequestrin, cardiac and slow skeletal muscle - northern leopard frog (fragment)
 N:Alternate names: 58K dihydropyridine-binding protein; aspartactin; calmitine; laminin-
 C:Species: Rana pipiens (northern leopard frog)
 C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 16-Jul-1999
 C:Accession: A61230
 R:McLeod, A.G.; Shen, A.C.Y.; Campbell, K.P.; Michalak, M.; Jorgensen, A.O.
 Circ. Res. 69, 344-359, 1991
 A:Title: Frog cardiac calsequestrin. Identification, characterization, and subcellular distribution.

A:Reference number: A61230; MUID:91316784; PMID:1960177
 A:Accession: A61230
 A:Molecule type: protein
 A:Residues: 1-9 <MCL>

C:Comment: Calsequestrin is a high-capacity and moderate-affinity calcium binding protein.
 C:Comment: Calsequestrin acts as a calcium buffer, and the release of calcium bound to calsequestrin can be phosphorylated in vivo.
 C:Superfamily: calsequestrin
 C:Keywords: calcium binding; cardiac muscle; glycoprotein; heart; phosphoprotein; skeletal

Query Match 76.9%; Score 10; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 1 EE 2

RESULT 12

C24180
 fibrinogen beta chain - Japanese macaque (fragment)
 N:Contains: fibrinopeptide B
 C:Species: Macaca fuscata (Japanese macaque)
 C:Date: 05-Jun-1988 #sequence_revision 10-Mar-1994 #text_change 26-Jan-1996
 C:Accession: C24180
 R:Nakamura, S.; Takenaka, O.; Takahashi, K.
 J. Biochem. 97, 1487-1492, 1985
 A:Title: Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and patas monkey (Mandrillus ssp.)
 A:Reference number: A91990; MUID:85289140; PMID:3928610
 A:Accession: C24180
 A:Molecule type: protein
 A:Residues: 1-9 <NAK>
 C:Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen C

Query Match 76.9%; Score 10; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 13

D24180
 fibrinogen beta chain - red guenon (fragment)
 N:Contains: fibrinopeptide B
 C:Species: Erythrocebus patas (red guenon, hussar)
 C:Date: 05-Jun-1988 #sequence_revision 10-Mar-1994 #text_change 26-Jan-1996
 C:Accession: D24180
 R:Nakamura, S.; Takenaka, O.; Takahashi, K.
 J. Biochem. 97, 1487-1492, 1985
 A:Title: Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and patas monkey (Mandrillus ssp.)
 A:Reference number: A91990; MUID:85289140; PMID:3928610
 A:Accession: D24180
 A:Molecule type: protein
 A:Residues: 1-9 <NAK>
 C:Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen C

Query Match 76.9%; Score 10; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 14

C60070
 gastrin - domestic ferret (fragment)
 C:Species: Mustela putorius furo (domestic ferret)
 C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 18-Jun-1993
 C:Accession: C60070
 R:Desmond, H.; Varro, A.; Young, J.; Gregory, H.; Nemeth, J.; Dockray, G.J.
 Regul. Pept. 25, 223-233, 1989
 A:Title: The constitution and properties of phosphorylated and unphosphorylated C-terminal fragments of gastrin
 A:Reference number: A60070; MUID:89331947; PMID:2756156
 A:Accession: C60070

A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <DES>

Query Match 76.9%; Score 10; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 3 EE 4

RESULT 15

PH0108
late G1-69 protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH0108
R;Nikaido, T.; Bradley, D.W.; Pardee, A.B.
Exp. Cell Res. 192, 102-109, 1991
A;Title: Molecular cloning of transcripts that accumulate during the late G1 phase in cu
A;Reference number: PH0108; MUID:91079351; PMID:1984406
A;Accession: PH0108
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-9 <NIK>

Query Match 76.9%; Score 10; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
||
Db 7 EE 8

Search completed: February 11, 2003, 18:21:19
Job time : 6.38144 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:07:35 ; Search time 2.21649 Seconds
(without alignments)
93.563 Million cell updates/sec

Title: US-09-380-738a-5
Perfect score: 13
Sequence: 1 XXXEE 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	10	76.9	5	1 TRM3_ECOLI	P13973 escherichia
2	10	76.9	7	1 UF03_MOUSE	P38641 mus musculus
3	10	76.9	9	1 FIBB_ERIPA	P19346 erythrocebu
4	10	76.9	9	1 FIBB_MACFU	P19345 macaca fusc
5	10	76.9	10	1 COXA_ONCMY	P80328 oncorhynch
6	10	76.9	10	1 COXO_RAT	P80432 rattus norv
7	10	76.9	10	1 SPI_HAIRO	Q10997 halocynthia
8	10	76.9	11	1 COXA_CANFA	P99501 canis fami
9	10	76.9	12	1 CALM_TETTH	Q05055 tetrahymena
10	10	76.9	12	1 HS9A_RAT	P82995 rattus norv
11	10	76.9	12	1 PA2B_VIPBO	P31859 vipera beru
12	10	76.9	13	1 BP37_LEUMA	P81754 leucophaea
13	10	76.9	13	1 EI21_LITRU	P82097 litorea rub
14	10	76.9	13	1 EI22_LITRU	P82098 litorea rub
15	10	76.9	13	1 PEDI_HYDAT	P80578 hydra atten
16	10	76.9	13	1 UHAI_HUMAN	P40928 homo sapien
17	10	76.9	14	1 DHSI_ANACY	P17874 anabaena cy
18	10	76.9	14	1 FIBA_HORSE	P14452 equus cabal
19	10	76.9	14	1 FIBB_MANLE	P14474 mandrillus
20	10	76.9	14	1 NSK2_SARBU	P41493 sarcophaga
21	10	76.9	14	1 UC15_MAIZE	P80621 zea mays (m
22	10	76.9	15	1 ESTJ_MANSE	P19985 manduca sex
23	10	76.9	15	1 GR78_HORSE	P16392 equus cabal
24	10	76.9	15	1 METK_MAIZE	P80616 zea mays (m
25	10	76.9	15	1 UN01_PINPS	P81106 pinus pinas
26	10	76.9	15	1 UN04_PINPS	P81673 pinus pinas
27	10	76.9	16	1 ALLI_CALVO	P41839 calliphora
28	10	76.9	16	1 FIBA_EQUAS	P14449 equus asinu
29	10	76.9	16	1 IBP4_PIG	P24854 sus scrofa
30	10	76.9	16	1 UPAB_HUMAN	P31935 homo sapien
31	10	76.9	17	1 ALYS_MYCPH	P81528 mycobacteri
32	10	76.9	17	1 GAST_MACMU	P33714 macaca mula
33	10	76.9	17	1 TRP2_LEUMA	P81733 leucophaea

34 10 76.9 18 1 OBP_LYMDI P34173 lymantria d
35 10 76.9 19 1 FIBB_HORSE P14471 equus cabal
36 10 76.9 19 1 FIBB_LAMGL P14473 lama glama
37 10 76.9 19 1 FIBB_TAPTE P14539 tapirus ter
38 10 76.9 19 1 FIBB_VULVU P14482 vulpus vulp
39 10 76.9 19 1 OXLA_OPHHA P81383 ophiophagus
40 10 76.9 19 1 TPIS_CLOPA P81348 clostridium
41 10 76.9 20 1 BIP_PHAVU P80089 phascelus v
42 10 76.9 20 1 CATY_ACIRA P81422 acinetobact
43 10 76.9 20 1 COXA_THUOB P80972 thunnus obe
44 10 76.9 20 1 COXF_ONCMY P80329 oncorhynchu
45 10 76.9 20 1 CPXX_RHORH P31718 rhodococcus

ALIGNMENTS

RESULT 1
TRM3_ECOLI
ID TRM3_ECOLI STANDARD; PRT; 5 AA.
AC P13973:
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Tram protein (Fragment).
GN TRAM.
OS Escherichia coli.
OG Plasmid IncFII R100.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88227859; PubMed=2836369;
RA Inamoto S., Yoshioaka Y., Ohtsubo E.;
RT "Identification and characterization of the products from the traJ
and traY genes of plasmid R100.";
RL J. Bacteriol. 170:2749-2757(1988).
CC -!- FUNCTION: TRANSFER GENE PROTEIN. IS INVOLVED IN THE CONJUGATION
PROCESS OF BACTERIAL CELLS FOR THE EXCHANGE OF PLASMID DNA.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: TO TRAM PROTEIN OF OTHER PLASMIDS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

CC EMBL; M20941; -; NOT_ANNOTATED_CDS.
DR PIR; A32014; A32014.
KW Conjugation; Plasmid; DNA-binding.
FT NON_TER 1
SQ SEQUENCE 5 AA; 634 MW; 6B1B1AA443500000 CRC64;
Query Match 76.9%; Score 10; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
DB 4 EE 5
RESULT 2
UF03_MOUSE
ID UF03_MOUSE STANDARD; PRT; 7 AA.
AC P38641:
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)

DE Unknown protein from 2D-page of fibroblasts (P36) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Fibroblast;
 RX MEDLINE=95009907; PubMed=7523108;
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
 RT "Separation and sequencing of familial and novel murine proteins
 using preparative two-dimensional gel electrophoresis.";
 RL Electrophoresis 15:735-745(1994).
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.1, ITS MW IS: 36 kDa.
 FT NON_TER 7
 SQ SEQUENCE 7 AA; 842 MW; 6AA72BDD1B1180 CRC64;

Query Match 76.9%; Score 10; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 3
 FIBB_ERYPA STANDARD; PRT; 9 AA.
 ID FIBB_ERYPA
 AC P19346;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
 GN Erythrocebus patas (Red guenon) (Hussar).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Erythrocebus.
 OX NCBI_TaxID=9538;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85289140; PubMed=3928610;
 RA Nakamura S., Takenaka O., Takahashi K.;
 RT "Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and
 patas monkey (Erythrocebus patas): their amino acid sequences,
 restricted mutations, and a molecular phylogeny for macaques,
 guenons, and baboons.";
 RL J. Biochem. 97:1487-1492(1985).
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 CC PIR: D24180; D24180.
 DR InterPro: IPR002181; Fibrinogen_C.
 DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KW Blood coagulation; Plasma.
 FT PEPTIDE 1 9 FIBRINOPEPTIDE B.
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 1020 MW; 69FE7879C732CB1B CRC64;

Query Match 76.9%; Score 10; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 4
 FIBB_MACFU STANDARD; PRT; 9 AA.
 ID FIBB_MACFU
 AC P19345;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
 GN Macaca fuscata fuscata (Japanese macaque).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9543;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85289140; PubMed=3928610;
 RA Nakamura S., Takenaka O., Takahashi K.;
 RT "Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and
 patas monkey (Erythrocebus patas): their amino acid sequences,
 restricted mutations, and a molecular phylogeny for macaques,
 guenons, and baboons.";
 RL J. Biochem. 97:1487-1492(1985).
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 CC PIR: D24180; D24180.
 DR InterPro: IPR002181; Fibrinogen_C.
 DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KW Blood coagulation; Plasma.
 FT PEPTIDE 1 9 FIBRINOPEPTIDE B.
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 1020 MW; 69FE7879C732CB1B CRC64;

Query Match 76.9%; Score 10; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 2 EE 3

Db 2 EE 3
 RESULT 4
 FIBB_MACFU STANDARD; PRT; 9 AA.
 ID FIBB_MACFU
 AC P19345;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
 GN Macaca fuscata fuscata (Japanese macaque).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9543;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85289140; PubMed=3928610;
 RA Nakamura S., Takenaka O., Takahashi K.;
 RT "Fibrinopeptides A and B of Japanese monkey (Macaca fuscata) and
 patas monkey (Erythrocebus patas): their amino acid sequences,
 restricted mutations, and a molecular phylogeny for macaques,
 guenons, and baboons.";
 RL J. Biochem. 97:1487-1492(1985).
 CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 CC PIR: C24180; C24180.
 DR InterPro: IPR002181; Fibrinogen_C.
 DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KW Blood coagulation; Plasma.
 FT PEPTIDE 1 9 FIBRINOPEPTIDE B.
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 1038 MW; 69FE65B9C735BB1B CRC64;

Query Match 76.9%; Score 10; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 5
 COXA_ONCMY STANDARD; PRT; 10 AA.
 ID COXA_ONCMY
 AC P80328;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome c oxidase polypeptide Va (EC 1.9.3.1) (Fragment).
 OS Onchorynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94237150; PubMed=8181469;
 RA Freund R., Kadenbach B.;
 RT "Identification of tissue-specific isoforms for subunits Vb and VIIa
 of cytochrome c oxidase isolated from rainbow trout.";
 RL Eur. J. Biochem. 221:1111-1116(1994).

Query Match 76.9%; Score 10; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 5
 COXA_ONCMY STANDARD; PRT; 10 AA.
 ID COXA_ONCMY
 AC P80328;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome c oxidase polypeptide Va (EC 1.9.3.1) (Fragment).
 OS Onchorynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94237150; PubMed=8181469;
 RA Freund R., Kadenbach B.;
 RT "Identification of tissue-specific isoforms for subunits Vb and VIIa
 of cytochrome c oxidase isolated from rainbow trout.";
 RL Eur. J. Biochem. 221:1111-1116(1994).

Query Match 76.9%; Score 10; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 5
 COXA_ONCMY STANDARD; PRT; 10 AA.
 ID COXA_ONCMY
 AC P80328;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome c oxidase polypeptide Va (EC 1.9.3.1) (Fragment).
 OS Onchorynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94237150; PubMed=8181469;
 RA Freund R., Kadenbach B.;
 RT "Identification of tissue-specific isoforms for subunits Vb and VIIa
 of cytochrome c oxidase isolated from rainbow trout.";
 RL Eur. J. Biochem. 221:1111-1116(1994).

Query Match 76.9%; Score 10; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 5
 COXA_ONCMY STANDARD; PRT; 10 AA.
 ID COXA_ONCMY
 AC P80328;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome c oxidase polypeptide Va (EC 1.9.3.1) (Fragment).
 OS Onchorynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94237150; PubMed=8181469;
 RA Freund R., Kadenbach B.;
 RT "Identification of tissue-specific isoforms for subunits Vb and VIIa
 of cytochrome c oxidase isolated from rainbow trout.";
 RL Eur. J. Biochem. 221:1111-1116(1994).

Query Match 76.9%; Score 10; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 ||
 Db 2 EE 3

RESULT 5
 COXA_ONCMY STANDARD; PRT; 10 AA.
 ID COXA_ONCMY
 AC P80328;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome c oxidase polypeptide Va (EC 1.9.3.1) (Fragment).
 OS Onchorynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94237150; PubMed=8181469;
 RA Freund R., Kadenbach B.;
 RT "Identification of tissue-specific isoforms for subunits Vb and VIIa
 of cytochrome c oxidase isolated from rainbow trout.";
 RL Eur. J. Biochem. 221:1111-1116(1994).

CC -!- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C
 CC OXIDASE, THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.
 DR PIR; S43625; S43625.
 KW Oxidoreductase; Heme; Mitochondrion; Inner membrane.
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1144 MW; C535C5B1AB02C33D CRC64;
 Query Match 76.9%; Score 10; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 Db 9 EE 10
 RESULT 6
 COXO_RAT
 ID COXO_RAT STANDARD; PRT; 10 AA.
 AC P80432;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytochrome c oxidase polypeptide VIIC, mitochondrial (EC 1.9.3.1)
 DE (VIIIA) (Fragment).
 GN COX7C OR COX7CL.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE.
 RC STRAIN-Wistar; TISSUE=Liver, and Heart;
 RX MEDLINE=95324529; PubMed=7601105;
 RA Schaeffer H., Noack H., Halangk W., Brandt U., von Jagow G.;
 RT "Cytochrome-c oxidase in developing rat heart. Enzymic properties and
 RT amino-terminal sequences suggest identity of the fetal heart and the
 RT adult liver isoform."
 RL Eur. J. Biochem. 230:235-241(1995).
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE. THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIC FAMILY.
 KW Oxidoreductase; Mitochondrion.
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1117 MW; 126DE767687B1B0B CRC64;
 Query Match 76.9%; Score 10; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 Db 4 EE 5
 RESULT 7
 SPL_HALRO
 ID SPL_HALRO STANDARD; PRT; 10 AA.
 AC Q10997;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Serine proteinase inhibitor (Fragment).
 OS Halocynthia roretzi (Sea squirt).
 CC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 CC Stolidobranchia; Pyuridae; Halocynthia.

OX NCBI_TaxID=7729;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Hemolymph;
 RX MEDLINE=96321313; PubMed=8759295;
 RA Shishikura F., Abe T., Ohtake S.-I., Tanaka K.;
 RT "Purification and characterization of a 58,000-Da proteinase
 RT inhibitor from the hemolymph of a solitary ascidian, Halocynthia
 RT roretzi."
 RL Comp. Biochem. Physiol. 114B:1-9(1996).
 CC -!- FUNCTION: STRONGLY INHIBITS TRYPSIN AND PLASMA ENZYME(S) ACTIVITY.
 CC -!- SUBUNIT: MONOMER.
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
 DR InterPro; IPR000215; Serpin.
 DR PROSITE; PS00284; SERPIN; PARTIAL.
 KW Serpin; Serine protease inhibitor; Glycoprotein; Plasma.
 FT NON_TER 10 10
 SQ SEQUENCE 10 AA; 1104 MW; 4225C73B1B187AA3 CRC64;
 Query Match 76.9%; Score 10; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 Db 6 EE 7
 RESULT 8
 COXA_CANFA
 ID COXA_CANFA STANDARD; PRT; 11 AA.
 AC P99501;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cytochrome c oxidase polypeptide Va (EC 1.9.3.1) (Fragment).
 GN COX5A.
 OS Canis familiaris (Dog).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Heart;
 RX MEDLINE=98163340; PubMed=9504812;
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
 RT dog heart proteins."
 RL Electrophoresis 18:2795-2802(1997).
 CC -!- FUNCTION: THIS IS THE HEME A-CONTAINING CHAIN OF CYTOCHROME C
 CC OXIDASE. THE TERMINAL OXIDASE IN MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VA FAMILY.
 DR HSC-2DPAGE; P99501; DOG.
 DR InterPro; IPR003204; Cyt_c_ox5A.
 DR Pfam; PF02284; COX5A; 1.
 KW Oxidoreductase; Heme; Mitochondrion; Inner membrane.
 FT NON_TER 11 11
 SQ SEQUENCE 11 AA; 1274 MW; 910B35C5B1AB11F5 CRC64;
 Query Match 76.9%; Score 10; DB 1; Length 11;
 Best Local Similarity 100.0%; Pred. No. 8.5e+02;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 Db 9 EE 10
 RESULT 9
 CALM_TETH

ID CALM_TETTH STANDARD; PRT; 12 AA.
 AC Q0505;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Calmodulin (Fragment).
 OS Tetrahymena thermophila.
 CC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
 CC Tetrahymenina; Tetrahymena.
 CC NCBI_TaxID=5911;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93281388; PubMed=8506136;
 RA Katoh M., Hirono M., Takemasa T., Kimura M., Watanabe Y.;
 RT "A microcnucleus-specific sequence exists in the 5'-upstream region of
 RL nucleic acid gene in Tetrahymena thermophila.";
 RL Nucleic Acids Res. 21:2409-2414(1993).
 CC -1- FUNCTION: CALMODULIN MEDIATES THE CONTROL OF A LARGE NUMBER OF
 CC ENZYMES BY CA(++) . AMONG THE ENZYMES TO BE STIMULATED BY THE
 CC CALMODULIN-CA(++) COMPLEX ARE A NUMBER OF PROTEIN KINASES AND
 CC PHOSPHATASES.
 CC -1- MISCELLANEOUS: THIS PROTEIN HAS FOUR FUNCTIONAL CALCIUM-BINDING
 CC SITES.
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D12774; BAA02239.1; -;
 DR InterPro; IPR002048; EF-hand.
 DR PROSITE; PS00018; EF-HAND; PARTIAL.
 KW Calcium-binding; Repeat; Acetylation.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
 FT NON_TER 12 12
 SQ SEQUENCE 12 AA; 1393 MW; 83F31CD443DB1B01 CRC64;
 Query Match 76.9%; Score 10; DB 1; Length 12;
 Best Local Similarity 100.0%; Pred. No. 9.3e+02;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 DB 6 EE 7
 RESULT 10
 HS9A_RAT STANDARD; PRT; 12 AA.
 ID HS9A_RAT
 AC P82995;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Heat shock protein Hsp 90-alpha (Fragment).
 DE HSPCA.
 GN HSPCA.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=21589773; PubMed=11732320;
 RA Langer T., Rasold H.;
 RT "Isolation and quantification of the heat shock protein 90 alpha and
 RT beta isoforms from rat liver.";
 RL Protoplasma 218:54-56(2001).
 CC -1- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY

CC (BY SIMILARITY).
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
 DR InterPro; IPR001404; Hsp90.
 DR PROSITE; PS00298; HSP90; PARTIAL.
 KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
 FT MOD_RES 4 4
 FT MOD_RES 6 6 PHOSPHORYLATION (BY DS-DNA KINASE) (BY
 FT SIMILARITY).
 FT MOD_RES 12 12 PHOSPHORYLATION (BY DS-DNA KINASE) (BY
 FT SIMILARITY).
 SQ SEQUENCE 12 AA; 1432 MW; DE47C322CAB6C1B6 CRC64;
 Query Match 76.9%; Score 10; DB 1; Length 12;
 Best Local Similarity 100.0%; Pred. No. 9.3e+02;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 DB 2 EE 3
 RESULT 11
 PA2B_VIPBO STANDARD; PRT; 12 AA.
 ID PA2B_VIPBO
 AC P31859;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phospholipase A2, basic (EC 3.1.1.4) (Phosphatidylcholine
 DE 2-acylhydrolase) (Fragment).
 OS Vipera berus orientalis (Viper).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;
 CC Viperidae; Viperinae; Vipera.
 CC NCBI_TaxID=31157;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=74128698; PubMed=4206446;
 RA Delori P.J.;
 RT "Purification and physicochemical, chemical and biological properties
 RT of a toxic A2 phospholipase isolated from the venom of viperidae
 RT snakes: Vipera berus.";
 RL Biochimie 55:1031-1045(1973).
 CC -1- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
 CC 2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
 CC -1- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-
 CC acylglycerophosphocholine + a fatty acid anion.
 CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
 DR InterPro; IPR001211; PhospholipaseA2.
 DR PROSITE; PS00118; PA2_HIS; PARTIAL.
 DR PROSITE; PS00119; PA2_ASP; PARTIAL.
 KW Hydrolase; Lipid degradation; Calcium; Venom.
 FT NON_TER 12 12
 SQ SEQUENCE 12 AA; 1412 MW; D054351A89969879 CRC64;
 Query Match 76.9%; Score 10; DB 1; Length 12;
 Best Local Similarity 100.0%; Pred. No. 9.3e+02;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 EE 5
 DB 11 EE 12
 RESULT 12
 BP37_LEUMA STANDARD; PRT; 13 AA.
 ID BP37_LEUMA
 AC P81754;

```

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Brain protein 37F3.
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Brain;
RX MEDLINE=97269266; PubMed=9114447;
RA Muren J.E., Naessel D.R.;
RT "Seven tachykinin-related peptides isolated from the brain of the
RT Madeira cockroach; evidence for tissue-specific expression of
RT isoforms."
RL Peptides 18:7-15(1997).
CC -!- TISSUE SPECIFICITY: BRAIN.
CC -!- MASS SPECTROMETRY: MW=1435.7; METHOD=MALDI.
SQ SEQUENCE 13 AA; 1436 MW; 8E532CDE9A7D1B5 CRC64;

Query Match 76.9%; Score 10; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
DB 4 EE 5

RESULT 13
EI21_LITRU STANDARD; PRT; 13 AA.
AC P82097;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Electrin 2.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litori electrica. Comparison with the skin peptides from Litoria
RT rubella."
RL Aust. J. Chem. 52:639-645(1999).
CC -!- SUBCELLULAR LOCATION: SECRETED.
KW Amphibian skin; Amidation.
FT MOD_RES 13
SQ SEQUENCE 13 AA; 1599 MW; C1808EF326F57322 CRC64;

Query Match 76.9%; Score 10; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
DB 2 EE 3

RESULT 14
EI22_LITRU STANDARD; PRT; 13 AA.
AC P82098;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```

DE Electrin 2.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litori electrica. Comparison with the skin peptides from Litoria
RT rubella."
RL Aust. J. Chem. 52:639-645(1999).
CC -!- SUBCELLULAR LOCATION: SECRETED.
KW Amphibian skin; Amidation.
FT MOD_RES 13
SQ SEQUENCE 13 AA; 1598 MW; C1808EF33B357322 CRC64;

Query Match 76.9%; Score 10; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
DB 2 EE 3

RESULT 15
PEDI_HYDAT STANDARD; PRT; 13 AA.
AC P80578;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Pedin.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE.
RX MEDLINE=96232307; PubMed=8674432;
RA Hoffmeister S.A.H.;
RT "Isolation and characterization of two new morphogenetically active
RT peptides from Hydra vulgaris."
RL Development 122:1941-1948(1996).
CC -!- FUNCTION: MORPHOGENETICALLY ACTIVE PEPTIDE. ACTIVE IN FOOT
CC DEVELOPMENT.
KW Morphogen.
SQ SEQUENCE 13 AA; 1512 MW; 6F5266E3F37632CB CRC64;

Query Match 76.9%; Score 10; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
DB 1 EE 2

Search completed: February 11, 2003, 18:17:49
Job time : 3.21649 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2003, 18:13:41 ; Search time 5.61856 Seconds
(without alignments)
183.363 Million cell updates/sec

Title: US-09-380-738A-5
Perfect score: 13
Sequence: 1 XXXEE 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriopl:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	10	76.9	7	6	Q28742
2	10	76.9	8	8	Q8WFR5
3	10	76.9	8	8	Q8W8G6
4	10	76.9	8	8	Q8W8G5
5	10	76.9	8	8	Q8W8G4
6	10	76.9	8	8	Q8W8G3
7	10	76.9	8	8	Q8W8G2
8	10	76.9	8	10	Q42507
9	10	76.9	8	13	Q902V5
10	10	76.9	9	6	Q9TRW2
11	10	76.9	9	8	Q8WFT4
12	10	76.9	9	8	Q8WFS4
13	10	76.9	9	8	Q8W8X4
14	10	76.9	9	8	Q8W8W6
15	10	76.9	9	8	Q8W8W5
16	10	76.9	9	10	P82440

17	10	76.9	9	12	Q84333
18	10	76.9	10	2	P74843
19	10	76.9	10	4	Q9UCQ8
20	10	76.9	10	4	Q8WTT4
21	10	76.9	10	4	Q9UCQ4
22	10	76.9	10	5	P82384
23	10	76.9	10	6	Q9GKI5
24	10	76.9	10	6	Q9GKI4
25	10	76.9	10	8	Q8WFT6
26	10	76.9	10	8	Q8WFT5
27	10	76.9	10	11	Q9OVK8
28	10	76.9	10	11	Q9OVE8
29	10	76.9	10	11	Q9QVE7
30	10	76.9	10	11	Q9QVJ6
31	10	76.9	10	11	Q9QVJ5
32	10	76.9	10	12	Q83978
33	10	76.9	11	10	Q06626
34	10	76.9	11	12	Q84073
35	10	76.9	11	13	Q90735
36	10	76.9	11	15	P88018
37	10	76.9	12	2	Q9L4M9
38	10	76.9	12	2	Q9X6Y0
39	10	76.9	12	2	Q93U04
40	10	76.9	12	2	Q8VLX8
41	10	76.9	12	4	Q96PK0
42	10	76.9	12	4	Q9UC37
43	10	76.9	12	9	Q8SCJ2
44	10	76.9	12	10	P82328
45	10	76.9	12	10	P82342

ALIGNMENTS

RESULT 1

Q28742 PRELIMINARY; PRT; 7 AA.
 AC Q28742;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Alpha-myosin heavy chain (Fragment).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OX Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 RN NCBI_TaxID=9986;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84221901; PubMed=6328491;
 RA Friedman D.J., Umeda P.K., Sinha A.M., Hsu H.-J., Jokovic S.,
 RA Rabinowitz M.;
 RT "Characterization of genomic clones specifying rabbit alpha- and beta-
 RT ventricular myosin heavy chains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3044-3048(1984).
 DR EMBL; K01698; AAA31415.1; -;
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 916 MW; 6B1B1AA1E6932680 CRC64;

Query Match 76.9%; Score 10; DB 6; Length 7;
 Best Local Similarity 100.0%; Pred. No. 6.7e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EE 5
 Db 6 EE 7

RESULT 2

Q8WFR5 PRELIMINARY; PRT; 8 AA.
 ID Q8WFR5
 AC Q8WFR5;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)

```

DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN COII.
OS Diadema paucispinum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematacea; Diadematoida; Diadematidae;
OC Diadema.
OX NCBI_TaxID=145530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H11;
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RL phylogeography of the sea urchin Diadema.";
DR Evolution 55:955-975(2001).
DR EMBL; AY012959; AAL33852.1; -.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 954 MW; C41B173B46DDC2CE CRC64;

Query Match 76.9%; Score 10; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
Db 7 EE 8

RESULT 3
Q8W8G6 PRELIMINARY; PRT; 8 AA.
ID Q8W8G6;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DE Cytochrome oxidase subunit II (Fragment).
GN COII.
OS Diadema mexicanum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematacea; Diadematoida; Diadematidae;
OC Diadema.
OX NCBI_TaxID=105359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GLA121, AND GLA124;
RX MEDLINE=21561594; PubMed=11703875;
RA Lessios H.A., Garrido M.J., Kessing B.D.;
RT "Demographic history of Diadema antillarum, a keystone herbivore on
RL Caribbean reefs.";
DR Proc. R. Soc. Lond., B, Biol. Sci. 268:2347-2353(2001).
DR EMBL; AY012796; AAL33821.1; -.
DR EMBL; AY012799; AAL33822.1; -.
DR EMBL; AY012803; AAL33823.1; -.
DR EMBL; AY012804; AAL33824.1; -.
DR EMBL; AY012806; AAL33825.1; -.
DR EMBL; AY012852; AAL33826.1; -.
DR EMBL; AY012854; AAL33828.1; -.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 1037 MW; 701B173B46DDC2D3 CRC64;

Query Match 76.9%; Score 10; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
Db 7 EE 8

RESULT 4
Q8W8G5 PRELIMINARY; PRT; 8 AA.
ID Q8W8G5;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN COII.
OS Diadema antillarum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematacea; Diadematoida; Diadematidae;
OC Diadema.
OX NCBI_TaxID=105358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA9414, DA9441, DIANI25, DIANI126, DIANI131, CN2, AND CN4;
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RL phylogeography of the sea urchin Diadema.";
DR Evolution 55:955-975(2001).
DR EMBL; AY012796; AAL33821.1; -.
DR EMBL; AY012799; AAL33822.1; -.
DR EMBL; AY012803; AAL33823.1; -.
DR EMBL; AY012804; AAL33824.1; -.
DR EMBL; AY012806; AAL33825.1; -.
DR EMBL; AY012852; AAL33826.1; -.
DR EMBL; AY012854; AAL33828.1; -.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 1037 MW; 701B173B46DDC2D3 CRC64;

Query Match 76.9%; Score 10; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
Db 7 EE 8

RESULT 5
Q8W8G4 PRELIMINARY; PRT; 8 AA.
ID Q8W8G4;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN COII.
OS Diadema mexicanum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematacea; Diadematoida; Diadematidae;
OC Diadema.
OX NCBI_TaxID=105359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C15, D3, D5, D6, CG66, G123, DM1, DM3, AND DM71;

```

```

RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RL phylogeography of the sea urchin Diadema.";
RN Evolution 55:955-975(2001).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C15, D3, D5, D6, CC66, GL23, DM1, DM3, AND DM71;
RX MEDLINE=21561594; PubMed=11703875;
RA Lessios H.A., Garrido M.J., Kessing B.D.;
RT "Demographic history of Diadema antillarum, a keystone herbivore on
RN Caribbean reefs.";
RP Proc. R. Soc. Lond., B, Biol. Sci. 268:2347-2353(2001).
DR EMBL; AY012908; AAL33837.1; -
DR EMBL; AY012911; AAL33838.1; -
DR EMBL; AY012913; AAL33839.1; -
DR EMBL; AY012914; AAL33840.1; -
DR EMBL; AY012915; AAL33842.1; -
DR EMBL; AY012940; AAL33847.1; -
DR EMBL; AY012949; AAL33849.1; -
DR EMBL; AY012950; AAL33850.1; -
DR EMBL; AY012951; AAL33851.1; -
KW Mitochondrion. 1
FT NON_TER 1
SQ SEQUENCE 8 AA; 1037 MW; 701B173B46DDC2D3 CRC64;

Query Match 76.9%; Score 10; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
DB 7 EE 8

RESULT 6
Q8W8G3 PRELIMINARY; PRT; 8 AA.
ID Q8W8G3;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DE Cytochrome oxidase subunit II (fragment).
GN COII.
OS Diadema paucispinum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diademataceae; Diadematoidea; Diadematiidae;
OC Diadema.
OX NCBI_TaxID=145530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HI2, HI3, HI5, HI6, HI9, HI15, AND HI19;
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RL phylogeography of the sea urchin Diadema.";
RN Evolution 55:955-975(2001).
DR EMBL; AY012960; AAL33853.1; -
DR EMBL; AY012961; AAL33854.1; -
DR EMBL; AY012962; AAL33855.1; -
DR EMBL; AY012963; AAL33856.1; -
DR EMBL; AY012964; AAL33857.1; -
DR EMBL; AY012965; AAL33858.1; -
DR EMBL; AY012966; AAL33859.1; -
KW Mitochondrion. 1
FT NON_TER 1
SQ SEQUENCE 8 AA; 1037 MW; 701B173B46DDC2D3 CRC64;

Query Match 76.9%; Score 10; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
DB 7 EE 8

RESULT 7
Q8W8G2 PRELIMINARY; PRT; 8 AA.
ID Q8W8G2;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DE Cytochrome oxidase subunit II (fragment).
GN COII.
OS Diadema savignyi.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diademataceae; Diadematoidea; Diadematiidae;
OC Diadema.
OX NCBI_TaxID=105360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DP751, SA2, SA7, SA10, GSA1, GSA3, GSA4, DOK17, AND DOK105;
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RL phylogeography of the sea urchin Diadema.";
RN Evolution 55:955-975(2001).
DR EMBL; AY013085; AAL33860.1; -
DR EMBL; AY013080; AAL33861.1; -
DR EMBL; AY013083; AAL33862.1; -
DR EMBL; AY013086; AAL33863.1; -
DR EMBL; AY013088; AAL33864.1; -
DR EMBL; AY013090; AAL33865.1; -
DR EMBL; AY013091; AAL33866.1; -
DR EMBL; AY013102; AAL33867.1; -
DR EMBL; AY013103; AAL33868.1; -
DR EMBL; AY013103; AAL33869.1; -
KW Mitochondrion. 1
FT NON_TER 1
SQ SEQUENCE 8 AA; 1037 MW; 701B173B46DDC2D3 CRC64;

Query Match 76.9%; Score 10; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
DB 7 EE 8

RESULT 8
Q42507 PRELIMINARY; PRT; 8 AA.
ID Q42507;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE Heat shock protein (fragment).
GN HSP70C OR HSP70A OR HSP70B.
OS Triticum aestivum (wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV MUSTANG;
RX MEDLINE=96189275; PubMed=8605312;
RA Joshi C.P., Kumar S., Nguyen H.T.;
RT "Application of modified differential display technique for cloning
RN and sequencing of the 3' region from three putative members of wheat
HSP70 gene family.";
RN Plant Mol. Biol. 30:641-646(1996).

```

DR EMBL; L41507; AAB02333.1; -;
 DR EMBL; L41505; AAB02331.1; -;
 DR EMBL; L41506; AAB02332.1; -;
 KW Heat shock.
 FT NON_TER
 SQ SEQUENCE 8 AA; 886 MW; 71B2CB1B10532768 CRC64;

Query Match 76.9%; Score 10; DB 10; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6.7e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 DB 5 EE 6

RESULT 9

Q902V5 PRELIMINARY; PRT; 8 AA.
 AC Q902V5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Adenylate kinase (Fragment).
 OS Fulica leucopetra.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Gruiformes; Rallidae; Fulica.
 OX NCBI_TaxID=156758;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shapero L.H., Dumbacher J.P.;
 RT "Adenylate kinase intron 5: a new nuclear locus for avian
 RT systematics";
 RL Auk 118:248-255(2001).
 DR EMBL; AF307898; AAK43537.1; -;
 KW Kinase
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 8 AA; 994 MW; 96333B19CB1B1866 CRC64;

Query Match 76.9%; Score 10; DB 13; Length 8;
 Best Local Similarity 100.0%; Pred. No. 6.7e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 DB 3 EE 4

RESULT 10

Q9TRW2 PRELIMINARY; PRT; 9 AA.
 AC Q9TRW2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE CALDESMON-PHOSPHORYLATION site (fragment).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=91378498; PubMed=1898046;
 RA Ikebe M., Hornick T.;
 RT "Determination of the phosphorylation sites of smooth muscle caldesmon
 RT by protein kinase C";
 RL Arch. Biochem. Biophys. 288:538-542(1991).
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 9 AA; 1018 MW; 8C901B10533735A5 CRC64;
 Query Match 76.9%; Score 10; DB 6; Length 9;

Best Local Similarity 100.0%; Pred. No. 6.7e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 DB 7 EE 8

RESULT 11

Q8WFT4 PRELIMINARY; PRT; 9 AA.
 AC Q8WFT4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Cytochrome oxidase subunit II (Fragment).
 GN COII.
 OS Diadema antillarum.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Diadematoidea; Diadematoidea; Diadematoidea;
 OC Diadema.
 OX NCBI_TaxID=105358;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DCA3;
 RX MEDLINE=21323357; PubMed=11430656;
 RA Lessios H.A., Kessing B.D., Pearse J.S.;
 RT "Population structure and speciation in tropical seas: global
 RT phylogeography of the sea urchin *Diadema*";
 RL Evolution 55:955-975(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DCA3;
 RX MEDLINE=21561594; PubMed=11703875;
 RA Lessios H.A., Garrido M.J., Kessing B.D.;
 RT "Demographic history of *Diadema antillarum*, a keystone herbivore on
 RT Caribbean reefs";
 RL Proc. R. Soc. Lond., B, Biol. Sci. 268:2347-2353(2001).
 DR EMBL; AY012858; AAL33832.2; -;
 KW Mitochondrion.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 9 AA; 1115 MW; EDF8DB1B173B46CA CRC64;

Query Match 76.9%; Score 10; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 6.7e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
 DB 6 EE 7

RESULT 12

Q8WFS4 PRELIMINARY; PRT; 9 AA.
 AC Q8WFS4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Cytochrome oxidase subunit II (Fragment).
 GN COII.
 OS Diadema mexicanum.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinoidea; Diadematoidea; Diadematoidea; Diadematoidea;
 OC Diadema.
 OX NCBI_TaxID=105359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GI25;
 RX MEDLINE=21323357; PubMed=11430656;
 RA Lessios H.A., Kessing B.D., Pearse J.S.;


```

RT "Population structure and speciation in tropical seas: global
RT phylogeography of the sea urchin Diadema.";
RL Evolution 55:955-975(2001).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=CC70;
RX MEDLINE=21561594; PubMed=11703875;
RA Lessios H.A., Garrido M.J., Kessing B.D.;
RT "Demographic history of Diadema antillarum, a keystone herbivore on
RT Caribbean reefs.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 268:2347-2353(2001).
DR EMBL; AY012941; AAL33843.1; -.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1151 MW; 2CED173B46DDC2D3 CRC64;

Query Match 76.9%; Score 10; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
Db 8 EE 9

RESULT 13
Q8W8X4 PRELIMINARY; PRT; 9 AA.
ID Q8W8X4
AC Q8W8X4;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN COII.
OS Diadema mexicanum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematacea; Diadematoidea; Diadematiidae;
OC Diadema.
OX NCBI_TaxID=105359;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CC70, AND CC117;
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RT phylogeography of the sea urchin Diadema.";
RL Evolution 55:955-975(2001).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=CC70, AND CC117;
RX MEDLINE=21561594; PubMed=11703875;
RA Lessios H.A., Garrido M.J., Kessing B.D.;
RT "Demographic history of Diadema antillarum, a keystone herbivore on
RT Caribbean reefs.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 268:2347-2353(2001).
DR EMBL; AY012920; AAL33843.1; -.
DR EMBL; AY012921; AAL33844.1; -.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1174 MW; 2B73173B46DDC2D3 CRC64;

Query Match 76.9%; Score 10; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
Db 8 EE 9

RESULT 14
Q8W8W6 PRELIMINARY; PRT; 9 AA.
ID Q8W8W6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN COII.
OS Diadema antillarum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematacea; Diadematoidea; Diadematiidae;
OC Diadema.
OX NCBI_TaxID=105358;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CN3, AND CN5;
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RT phylogeography of the sea urchin Diadema.";
RL Evolution 55:955-975(2001).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=CN3, AND CN5;
RX MEDLINE=21561594; PubMed=11703875;
RA Lessios H.A., Garrido M.J., Kessing B.D.;
RT "Demographic history of Diadema antillarum, a keystone herbivore on
RT Caribbean reefs.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 268:2347-2353(2001).
DR EMBL; AY012853; AAL33827.1; -.
DR EMBL; AY012855; AAL33829.1; -.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 9 AA; 1151 MW; 2CED173B46DDC2D3 CRC64;

Query Match 76.9%; Score 10; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5
Db 8 EE 9

RESULT 15
Q8W8W5 PRELIMINARY; PRT; 9 AA.
ID Q8W8W5;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Cytochrome oxidase subunit II (Fragment).
GN COII.
OS Diadema setosum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Diadematacea; Diadematoidea; Diadematiidae;
OC Diadema.
OX NCBI_TaxID=31175;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=DEL1, DEL2, AND DEL3, AND DEL5;
RX MEDLINE=21323357; PubMed=11430656;
RA Lessios H.A., Kessing B.D., Pearse J.S.;
RT "Population structure and speciation in tropical seas: global
RT phylogeography of the sea urchin Diadema.";
RL Evolution 55:955-975(2001).
DR EMBL; AY012890; AAL33833.1; -.
DR EMBL; AY012891; AAL33834.1; -.
DR EMBL; AY012892; AAL33835.1; -.
DR EMBL; AY012893; AAL33836.1; -.
KW Mitochondrion.
FT NON_TER 1

```

SQ SEQUENCE 9 AA; 1151 MW; 2CED173B46DDC2D3 CRC64;

Query Match 76.9%; Score 10; DB 8; Length 9;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EE 5

Db 8 EE 9

Search completed: February 11, 2003, 18:19:47
Job time : 6.61856 secs